



# STIC Search Report

## Biotech-Chem Library

STIC Database Tracking Number: 161820

TO: Celine Qian  
Location: REM-2A64&2C70  
Art Unit: 1636  
Thursday, August 11, 2005  
  
Case Serial Number: 09/913878

From: Edward Hart  
Location: Biotech-Chem Library  
REM-1A55  
Phone: 571-272-2512  
  
edward.hart@uspto.gov

### Search Notes

Examiner Qian,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart

**This Page Blank (uspto)**

1011N: EN Hart.

Access DB#

161820

## SEARCH REQUEST FORM

Scientific and Technical Information Center

Requester's Full Name: 09/913878 Examiner #: 78710 Date: 8/8/05  
 Art Unit: 1636 Phone Number: 2-0777 Serial Number: 09/913878  
 Mail Box and Bldg/Room Location: 2464 Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of invention: Isolation and characterization of an Crassa Silencing gene & uses thereof

Inventors (please provide full names):

Giuseppe et al.

Earliest Priority Filing Date: 2/22/1999

CRTE

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

8045 NA

Please search SEQ ID NO:1 from 2447-6652  
(commercial & interference database)

## STAFF USE ONLY

Type of Search		Vendors and cost where applicable
Searcher: _____	NA Sequence (#) <u>1</u>	STN _____
Searcher Phone #: _____	AA Sequence (#) _____	Dialog _____
Searcher Location: _____	Structure (#) _____	Questel/Orbit _____
Date Searcher Picked Up: <u>8/9/05</u>	Bibliographic _____	Dr. Link _____
Date Completed: <u>8/11/05</u>	Litigation _____	Lexis/Nexis _____
Searcher Prep. Review Time _____	Fulltext _____	Sequence Systems <u>04</u>
Clerical Prep. Time _____	Patent Family _____	WWW/Internet _____
Online Time _____	Other _____	Other (specify) _____

**This Page Blank (uspto)**



# STIC SEARCH RESULTS FEEDBACK FORM

## Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher* or contact:

Mary Hale, Information Branch Supervisor  
Remsen Bldg. 01 D86  
571-272-2507

## Voluntary Results Feedback Form

➤ I am an examiner in Workgroup:  Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature  
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC-Biotech-Chem Library Remsen Bldg.



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2005, 13:42:12 ; Search time 17518 Seconds  
(without alignments)  
11633.899 Million cell updates/sec

Title: US-09-913-878A-1\_COPY\_2447\_6652

Perfect score: 4206  
Sequence: 1 atgaaccttactactctag.....tcacagsgaatgagcatat 4206

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl: \*  
1: gb\_ba: \*  
2: gb\_bkg: \*  
3: gb\_in: \*  
4: gb\_on: \*  
5: gb\_ov: \*  
6: gb\_pat: \*  
7: gb\_ph: \*  
8: gb\_pl: \*  
9: gb\_pr: \*  
10: gb\_ro: \*  
11: gb\_sls: \*  
12: gb\_sy: \*  
13: gb\_un: \*  
14: gb\_vl: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4206	100.0	4206	8 NCR133528	AJ133528 Neurospor
2	4206	100.0	8045	6 AX034425	AX034425 Sequence
3	4202.8	99.9	4206	6 CQ803434	CQ803434 Sequence
4	3078	73.2	3078	6 CQ803436	CQ803436 Sequence
5	55.2	1.3	125020	9 AF429315	AF429315 Homo sapi
6	53.8	1.3	6688	8 AF468822	AF468822 Diaporthe
7	52.4	1.2	2000	6 AX655393	AX655393 Sequence
8	51.8	1.2	2000	6 AX655393	AX655393 Sequence
9	51.6	1.2	1164	6 A10243	A10243 S. murinus D
10	51.4	1.2	3524	8 AF443073	AF443073 Phomopsis
11	50.8	1.2	125020	9 AF429315	AF429315 Homo sapi
12	50	1.2	1610	1 AB038265	AB038265 Streptomy
13	50	1.2	1952	1 AF170068	AF170068 Streptomy
14	49	1.2	3731	6 ARI45904	ARI45904 Sequence
15	49	1.2	3731	8 LERDRP	Y10403 L. esculentu
16	48.6	1.2	11492	1 AB012149	AB012149 Xanthomon
17	48.4	1.2	1551	1 SSPY1518	Y15518 Streptomyce
18	48.4	1.2	1750	1 AF257177	AF257177 Streptomy
19	47	1.1	19601	1 AB070944	AB070944 Streptomy

20	47	1.1	299425	1 AP005049	AP005049 Streptomy
21	46.8	1.1	3674	8 AY574374	AY574374 Nicotiana
22	45.8	1.1	3800	1 SAR416377	SA416377 Streptomy
23	45.8	1.1	6204	8 AF411016	AF411016 Neurospor
24	45.6	1.1	2710	6 AX059555	AX059555 Sequence
25	45.4	1.1	6987	6 AY049072	AY049072 Diaporthe
26	44.6	1.1	194083	10 AC132253	AC132253 Mus muscu
27	44.6	1.1	349742	1 BX572090	BX572090 Prochloro
28	44.2	1.1	162959	2 AC019255	AC019255 Homo sapi
29	44.2	1.1	178217	9 AC022080	AC022080 Homo sapi
30	44.2	1.1	197866	2 AC068293	AC068293 Homo sapi
31	44	1.0	720	6 CQ718786	CQ718786 Sequence
32	44	1.0	1318	6 AF025354	AF025354 Homo sapi
33	44	1.0	1831	6 AX405860	AX405860 Sequence
34	44	1.0	1919	6 BC017798	BC017798 Homo sapi
35	44	1.0	1959	6 CQ492849	CQ492849 Sequence
36	44	1.0	2149	9 AK057696	AK057696 Homo sapi
37	44	1.0	75440	9 AP000786	AP000786 Homo sapi
38	44	1.0	118218	2 AP000670	AP000670 Homo sapi
39	44	1.0	190739	2 AP001361	AP001361 Homo sapi
40	44	1.0	192229	9 AP000943	AP000943 Homo sapi
41	43.8	1.0	34445	9 AC006291	AC006291 Homo sapi
42	43.8	1.0	75214	8 BX842597	BX842597 Neurospor
43	43.8	1.0	77292	8 BX842596	BX842596 Neurospor
44	43.8	1.0	114918	9 AC005188	AC005188 Homo sapi
45	43.8	1.0	153667	9 AC109439	AC109439 Homo sapi

#### ALIGNMENTS

RESULT 1					
LOCUS	NCR133528				
DEFINITION	Neurospora crassa qde-1 gene, partial.	4206 bp	DNA	linear	PLN 10-MAY-1999
ACCESSION	AJ133528				
VERSION	AJ133528.1 GI:4803726				
KEYWORDS	qde-1 gene; RNA-dependent RNA polymerase.				
SOURCE	Neurospora crassa				
ORGANISM	Neurospora crassa				
REFERENCE	1				
AUTHORS	Cogoni, C. and Macino, G.				
TITLE	Gene silencing in neurospora crassa requires a protein homologous				
JOURNAL	to RNA-dependent RNA polymerase				
REFERENCE	Unpublished				
AUTHORS	2 (bases 1 to 4206)				
TITLE	Cogoni, C.				
JOURNAL	Direct Submission				
	Submitted (08-MAR-1999) Cogoni C., Dept. Biotechnologie Cellulairi ed				
	Ematologia, Universita di Roma La Sapienza, Viale Regina Elena				
	324, 00161 Rome, ITALY				
FEATURES	Location/Qualifiers				
source	1..4206				
	/organism="Neurospora crassa"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:5141"				
gene	1..4206				
	/gene="qde-1"				
	1..4206				
	/gene="qde-1"				
	/function="gene silencing"				
	/note="putative"				
	/codon_start=1				
	/product="RNA-dependent RNA polymerase"				
	/protein_id="CA842634.1"				
	/db_xref="GI:4803727"				
	/db_xref="GOA:Q9Y7G6"				
	/db_xref="UniProt/TREMBL:Q9Y7G6"				
	/translation="MNPITPRKNSPVYEIINRLANDYNGLCQVADTTLTPRRKSL				
	AESEDDGRDKIKYRALNPLVWRKDSLNOAEANFFLEAKASSNWPKAHADDTLP				
	WSKEPPRAATAGQOMALQTVILLEVLRNFMPPNNPTPRTTGRITLSPSGLSRPTSTWT				

KRDEPANTVPADPRKSLTRSATGPPIHGAALPLKPPDVNTGSKPSELENLNC  
TKAKGLSDNVAAAAPVPIASALDKVPRIRANTRDPLTAGHRADVDSDPTDQ  
GTSGSVSFACRNNOSTOSFEAPSOPEKRPVADATVEFAPLHIESPCKTTS  
HINDOPLSSSQGETSFSTYSESPSSGSGALPEPSRNGLAESASQVHAP  
VVARLNIWPKPKMLHEAPLAIVAEVTLFPMCKVDLDESLGLYDSWSTARY  
TDIMKTLXRLDAPFGKPPKPPKPPVDFVTMPSPTSSVPPVYSKQDABEVIO  
PLIVYKPLPLMEFGKCLTRFRPDRFELLIPSTSPSSVPPVYSKQDABEVIO  
MLTWGSHLVGRQWRAFPAKDRGRKRLREFQLAEDPKI1KERVHFPAETITFRP  
DVEKTRSVPAEEVEORTBFRKVSQMLDMLQJDNMTWPHKLFR1QJLSTVAI  
MTLEPHQIRHKTDLSPSGTEVMDVGRMSVAKR1RDVLGLDSDVASQVRG  
SAKMWIVDDVDDDEDMIEITYPSORKECDPVDKHORTLEVESVABELASQNLQJ  
LPLVEDBARDKVXRQALIGRLINDLROSEBOKHALNREVPORWYBESRRATV  
SHGRVPLIAGLPDSQETLPLMNSGDPKQKTLQDIANDLQKRDITLKSXNITV  
GRSAV1YMIADPFWGLVEENHVHVSSEKFEDEESFTLLSDVDLVARSPAHFPSDIO  
RVRAVFPRELHSLKDV11PSTGDPVPLAKLSGGDYGDMAWCMDEI1VDGFVNAEM  
PLBPLSLRYLKQDITFKOLMAHSHTGSAKEQTYMIQSFHFLQJPNFLGCTNY  
KERLCYINNVSUNKPAILISLVNLDVSGKOGIVENEAQOALRELLGLGSLTPP  
MYKSDMLGEGEPHIIIDYLKFSIARPAIDIKELPAFNMAKKAADNDSPHMDPOLA  
SYTTFPEKISDKSRSSALLFTYTLKNR1GVBKEK1GR1VKNKEMRSDOP1PVRVNOY  
EKWCALITPEAMDSGANYSKVLRLLESLADEEMTMA1LRASTAFKLYYHSPKX  
VWQMAQRLAV1KAQMTSRPEGA.PALMTAFMYAGLMPDKKFTQYVARLEGGSEYP  
DPEYVEVLGDDDFDGLGFTGNDY"

## ORIGIN

Query Match	100.0%	Score 4206	DB 8	Length 4206
Best Local Similarity	100.0%	Pred. No. 0		
Matches 4206	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY 1	ATGAACCTTATTA	CTCCTAGAGAGAGAAATGACCCCGTCGAGAAATCATTAACCGGCTC	60	
Db 1	ATGAACCTTATTA	CTCCTAGAGAGAGAAATGACCCCGTCGAGAAATCATTAACCGGCTC	60	
QY 61	AATAAGACTACAA	CTTGAGGCTTCAGTGTGCGACAGACAACTTCACCCCACGCG	120	
Db 61	AATAAGACTACAA	CTTGAGGCTTCAGTGTGCGACAGACAACTTCACCCCACGCG	120	
QY 121	CGGAAGAGCTGCG	CGGAGAGTGAAGAGATTTCGGTGGCCATGACAAAGATCTACAGAGCC	180	
Db 121	CGGAAGAGCTGCG	CGGAGAGTGAAGAGATTTCGGTGGCCATGACAAAGATCTACAGAGCC	180	
QY 181	CTGAACCTTTCT	CTACTGCGGAGAGATGACTCCCTGAACAGGAGAGCAACTTCTTC	240	
Db 181	CTGAACCTTTCT	CTACTGCGGAGAGATGACTCCCTGAACAGGAGAGCAACTTCTTC	240	
QY 241	ATCGAGGCCAAAG	CTGCGAGCTGCAACTGAGTGCCTGAAGGCCAGCGCAGCTTACAGC	300	
Db 241	ATCGAGGCCAAAG	CTGCGAGCTGCAACTGAGTGCCTGAAGGCCAGCGCAGCTTACAGC	300	
QY 301	CTTCCGCTGCTCA	AGGAGAACTCCCGCGCGCTACTGCGCGCCAAACAATGGCATTGACG	360	
Db 301	CTTCCGCTGCTCA	AGGAGAACTCCCGCGCGCTACTGCGCGCCAAACAATGGCATTGACG	360	
QY 361	ACTGTGTGCTCG	AGTGTGCTTAATAGGTTATGCGACCTCCCAATAACACACAGAGTGA	420	
Db 361	ACTGTGTGCTCG	AGTGTGCTTAATAGGTTATGCGACCTCCCAATAACACACAGAGTGA	420	
QY 421	ACGTTTGAGAACT	TAAGGCGCCAAAGTGCCTGAAGGCCAGCCTCTACCAACACC	480	
Db 421	ACGTTTGAGAACT	TAAGGCGCCAAAGTGCCTGAAGGCCAGCCTCTACCAACACC	480	
QY 481	AAAGCGAAGATG	AGCCCGCAATGTCACTTTCGCTGATCGCGCAAAAGCGTGTGACT	540	
Db 481	AAAGCGAAGATG	AGCCCGCAATGTCACTTTCGCTGATCGCGCAAAAGCGTGTGACT	540	
QY 541	CGCTCTGCAAGG	TCTCTATTCAGCGCGCGCGGAGTACCCCTAAAGTTCCCGGATCCA	600	
Db 541	CGCTCTGCAAGG	TCTCTATTCAGCGCGCGCGGAGTACCCCTAAAGTTCCCGGATCCA	600	
QY 601	GTCGAATACCG	GTTCGAACGACCATCTCTCGAGATGAGATCTCAATCAGTGCACCAAG	660	
Db 601	GTCGAATACCG	GTTCGAACGACCATCTCTCGAGATGAGATCTCAATCAGTGCACCAAG	660	
QY 661	CGGCGCAAGGCA	AGTGTCTGATTAATGTTGCCGCTCGCGCGCGCGCGCGCGCTTATT	720	

Db 661	CGGCGCAAGGCA	AGTGTCTGATTAATGTTGCCGCTCGCGCGCGCGCGCGCGCTTATT	720
QY 721	GGCAGCGCTTGA	GACAAGGATACCGACTCGAAGGATGCCAATGAGAGATGCCACGCG	780
Db 721	GGCAGCGCTTGA	GACAAGGATACCGACTCGAAGGATGCCAATGAGAGATGCCACGCG	780
QY 781	ACAGGTCATAG	ACAGCGGACAGGATGATTCCTTTGATCATCTCAAGGCACTTCTAT	840
Db 781	ACAGGTCATAG	ACAGCGGACAGGATGATTCCTTTGATCATCTCAAGGCACTTCTAT	840
QY 841	GGTTCCAGTGT	CTTCAGCGCTTCCCGTCCATACAGACCTACCCAGATACCTTTTGA	900
Db 841	GGTTCCAGTGT	CTTCAGCGCTTCCCGTCCATACAGACCTACCCAGATACCTTTTGA	900
QY 901	GCTCTCTCTT	CAAGCCCAAGAGAGAGGCGCTGTGATGTCACCGCTTTGAGCTGGA	960
Db 901	GCTCTCTCTT	CAAGCCCAAGAGAGAGGCGCTGTGATGTCACCGCTTTGAGCTGGA	960
QY 961	CACCTGATTA	GATCTCTAGCAAGGAAAGAAACAACAGTCCCAATGATTAACAAGCC	1020
Db 961	CACCTGATTA	GATCTCTAGCAAGGAAAGAAACAACAGTCCCAATGATTAACAAGCC	1020
QY 1021	CTTTCACTGCT	CTCCAGGCTGAAACTTGTGAGACTTAATGATGCTTTCCAACT	1080
Db 1021	CTTTCACTGCT	CTCCAGGCTGAAACTTGTGAGACTTAATGATGCTTTCCAACT	1080
QY 1081	TCCGCGCGCA	GGGCGCAATTCGCGAGCGAGTCTCAAAATGACCTGCGTGGAGCGAA	1140
Db 1081	TCCGCGCGCA	GGGCGCGCAATTCGCGAGCGAGTCTCAAAATGACCTGCGTGGAGCGAA	1140
QY 1141	GAAGGCTGAT	CTCAGGTTCAAGTTCATGCCGCGGTGTCAGCTCGGCTGAGAAAT	1200
Db 1141	GAAGGCTGAT	CTCAGGTTCAAGTTCATGCCGCGGTGTCAGCTCGGCTGAGAAAT	1200
QY 1201	ATTTGGCCGAA	ATTTCCCAATGCTACAGAAAGCTCTCTGCTGTGATGGAAGTT	1260
Db 1201	ATTTGGCCGAA	ATTTCCCAATGCTACAGAAAGCTCTCTGCTGTGATGGAAGTT	1260
QY 1261	ACGAGACTTTA	TGACATGCAAGTGAACCTTGGAAGAAGAGGCTGGGCTTAAAGTAC	1320
Db 1261	ACGAGACTTTA	TGACATGCAAGTGAACCTTGGAAGAAGAGGCTGGGCTTAAAGTAC	1320
QY 1321	GACCTTCTG	CTGCTACCGCGCGATGTCACAGATATCTGGAAGCTCTTACCGGCTT	1380
Db 1321	GACCTTCTG	CTGCTACCGCGCGATGTCACAGATATCTGGAAGCTCTTACCGGCTT	1380
QY 1381	GATGCTTCT	CGTGTAAACCTTTCCAGAAAGCGCGCCAAAGAGCGTGTGCGACGCA	1440
Db 1381	GATGCTTCT	CGTGTGTAAACCTTTCCAGAAAGCGCGCCAAAGAGCGTGTGCGACGCA	1440
QY 1441	ATGACGGGCA	ACTTTGAGAGCAAGTGAAGCGGTGTTCTCTGCTGTTACAGACTAC	1500
Db 1441	ATGACGGGCA	ACTTTGAGAGCAAGTGAAGCGGTGTTCTCTGCTGTTACAGACTAC	1500
QY 1501	AATCCGACA	CACTGCTACTGCGCCCTTTTACTGTGAAGCTGAAGCGCTCATGTTT	1560
Db 1501	AATCCGACA	CACTGCTACTGCGCCCTTTTACTGTGAAGCTGAAGCGCTCATGTTT	1560
QY 1561	GAGCAGGCTG	TGATCAACCCGTCGGTTCGATGAGTGTTCAGAGATCTTTATA	1620
Db 1561	GAGCAGGCTG	TGATCAACCCGTCGGTTCGATGAGTGTTCAGAGATCTTTATA	1620
QY 1621	CGCTGCTT	CAAGCCCAAGCTGTAACCGCGGTGTGAGAAACAACAGGTCG	1680
Db 1621	CGCTGCTT	CAAGCCCAAGCTGTAACCGCGGTGTGAGAAACAACAGGTCG	1680
QY 1681	GTCGAAGATCA	TCAGTGTCTCACGATGGGCAACATTTCTTGTAGCGCGCAATGG	1740
Db 1681	GTCGAAGATCA	TCAGTGTCTCACGATGGGCAACATTTCTTGTAGCGCGCAATGG	1740
QY 1741	CGGCTTTCT	TGCGCAAGTGCAGTAACGAAACCTCTCAGGAGTTCCAGCTCCGC	1800
Db 1741	CGGCTTTCT	TGCGCAAGTGCAGTAACGAAACCTCTCAGGAGTTCCAGCTCCGC	1800



Dh 1741 CGCGCTTTCTTCGCGAAGATGCCGATACAGAAACCTCTCAGGAGTTCCAGCTCCG 1800  
Qy 1801 GCGGAGACCGGAACCATCATCAAGAGAGAGTCCACTTCTTGCGAGACCGGCATT 1860  
Db 1801 GCGGAGACCGGAACCATCATCAAGAGAGAGTCCACTTCTTGCGAGACCGGCATT 1860  
Qy 1861 ACGTTCGACCTGATGTGTTCAAGACAGAGATCTGCTCCGCGAGAGAACTGTAGAG 1920  
Db 1861 ACGTTCGACCTGATGTGTTCAAGACAGAGATCTGCTCCGCGAGAGAACTGTAGAG 1920  
Qy 1921 CAACGACCGAGTTCAAAAGTTAGTCAAAATGCTGACGCTGCTCTGCAACTGCAACAAC 1980  
Db 1921 CAACGACCGAGTTCAAAAGTTAGTCAAAATGCTGACGCTGCTCTGCAACTGCAACAAC 1980  
Qy 1981 ACTTGAGCGCGACCTCAAGTTGTCTCCGATTCAGAGCTGGTCTGAGTAAGACATAT 2040  
Db 1981 ACTTGAGCGCGACCTCAAGTTGTCTCCGATTCAGAGCTGGTCTGAGTAAGACATAT 2040  
Qy 2041 GCCATTATGACATTGGAGCCTCACCAAGTCAAGACCAAGACCGATCTTTTCACT 2100  
Db 2041 GCCATTATGACATTGGAGCCTCACCAAGTCAAGACCAAGACCGATCTTTTCACT 2100  
Qy 2101 TCAGGCACTGGCGAAGTATGATGATGACGCTGAGCGGCATGTCGCAAGCGTGGCAAG 2160  
Db 2101 TCAGGCACTGGCGAAGTATGATGATGACGCTGAGCGGCATGTCGCAAGCGTGGCAAG 2160  
Qy 2161 AGGATAGCGCATGTTCTCGGTTTGGGTGATGCTGCGCTCTGCTGCAAGGGCGGTTGGT 2220  
Db 2161 AGGATAGCGCATGTTCTCGGTTTGGGTGATGCTGCGCTCTGCTGCAAGGGCGGTTGGT 2220  
Qy 2221 TCGGCGCAAGGAAATGTGGTTATCGAGTTGACGACACAGCGATGAGATTTGATGAG 2280  
Db 2221 TCGGCGCAAGGAAATGTGGTTATCGAGTTGACGACACAGCGATGAGATTTGATGAG 2280  
Qy 2281 ACATACCCTGTCACGCCCAAGTGGAAATGCGACTTCTGTATTAATCAACGTACCTTC 2340  
Db 2281 ACATACCCTGTCACGCCCAAGTGGAAATGCGACTTCTGTATTAATCAACGTACCTTC 2340  
Qy 2341 GAAGTCGAGAGCGGCTTCTGAACTGAAAGTCACTGCTCTCAACCTACAGCTGTACT 2400  
Db 2341 GAAGTCGAGAGCGGCTTCTGAACTGAAAGTCACTGCTCTCAACCTACAGCTGTACT 2400  
Qy 2401 GTCTCGAAGATATAGGACGAGGACAAAGTGAAGATGCGCCAGCAATCGTGACCTT 2460  
Db 2401 GTCTCGAAGATATAGGACGAGGACAAAGTGAAGATGCGCCAGCAATCGTGACCTT 2460  
Qy 2461 ATCAACGATTTGCAACGACAGTTCAAGGCAAAAGCATGCTTGAATCGCCAGTGA 2520  
Db 2461 ATCAACGATTTGCAACGACAGTTCAAGGCAAAAGCATGCTTGAATCGCCAGTGA 2520  
Qy 2521 TTTTCGCAATGGGTTTTCGAGAGTTATTCAGTCCGCCAATCGAGTCAGCAAGGCGCT 2580  
Db 2521 TTTTCGCAATGGGTTTTCGAGAGTTATTCAGTCCGCCAATCGAGTCAGCAAGGCGCT 2580  
Qy 2581 GTGCTTTTCTTGCTGGGCTACTGACAGTCAAGAGAGACACTGAATCTTTGATGAC 2640  
Db 2581 GTGCTTTTCTTGCTGGGCTACTGACAGTCAAGAGAGACACTGAATCTTTGATGAC 2640  
Qy 2641 AGTGGGTTGATCCCAAGACAAAGTACTTGAAGACATCGCTGGGATCTTCAAAAG 2700  
Db 2641 AGTGGGTTGATCCCAAGACAAAGTACTTGAAGACATCGCTGGGATCTTCAAAAG 2700  
Qy 2701 CGGAAATGTGACAGTTGAAGTCCAGTGAACATCCGTCGCTGATGACATTAAT 2760  
Db 2701 CGGAAATGTGACAGTTGAAGTCCAGTGAACATCCGTCGCTGATGACATTAAT 2760  
Qy 2761 TACATGATGCGGATTTCTGGGGTGTGCTTGAAGAAATGAGTTCAATGTCGATTTCC 2820  
Db 2761 TACATGATGCGGATTTCTGGGGTGTGCTTGAAGAAATGAGTTCAATGTCGATTTCC 2820  
Qy 2821 TCAAAATTCAGAGACGAGAGAGATCTTTTACACTCTATCGACTGTATGTCCTGCTG 2880  
Db 2821 TCAAAATTCAGAGACGAGAGAGATCTTTTACACTCTATCGACTGTATGTCCTGCTG 2880

Qy 2881 GCGGATCCCCAGCCCATTTCCCTAGATATCCAAGGTTGAGCAGTCTTCAAGCA 2940  
Db 2881 GCGGATCCCCAGCCCATTTCCCTAGATATCCAAGGTTGAGCAGTCTTCAAGCA 2940  
Qy 2941 GAGCTCACAGTCTCAAGAGATGATATCTTCTTAATAAGAGATGATACCGTTGCT 3000  
Db 2941 GAGCTCACAGTCTCAAGAGATGATATCTTCTTAATAAGAGATGATACCGTTGCT 3000  
Qy 3001 AAGAAGTATCTGTGAGAGCTACGACGAGATATGGCCGTGCTGTGGATCCGAG 3060  
Db 3001 AAGAAGTATCTGTGAGAGCTACGACGAGATATGGCCGTGCTGTGGATCCGAG 3060  
Qy 3061 ATGTCGATGGTTTCTGCAATGCGAAATGCTCTGAGAGCCGACCTGTAGATCTTA 3120  
Db 3061 ATGTCGATGGTTTCTGCAATGCGAAATGCTCTGAGAGCCGACCTGTAGATCTTA 3120  
Qy 3121 AAGAAGCAAAACGACTTTCAACAACTTATGGCTCACACGGCACAGGCTCAGCGGC 3180  
Db 3121 AAGAAGCAAAACGACTTTCAACAACTTATGGCTCACACGGCACAGGCTCAGCGGC 3180  
Qy 3181 AAGAAGCAAGTACATACATATGATTCAGAAAGAGCTTCCATTTGCGCTCGACGCCAC 3240  
Db 3181 AAGAAGCAAGTACATACATATGATTCAGAAAGAGCTTCCATTTGCGCTCGACGCCAC 3240  
Qy 3241 TTTCTGGGCAATGCACTAACCTAACAAAGAGGCTGTATACATCAATAGTGTCT 3300  
Db 3241 TTTCTGGGCAATGCACTAACCTAACAAAGAGGCTGTATACATCAATAGTGTCT 3300  
Qy 3301 AACAGCGGCGCATCATTTCTTAAGTTCACTGTGGGAAACCTGTCGATCAGACAAAC 3360  
Db 3301 AACAGCGGCGCATCATTTCTTAAGTTCACTGTGGGAAACCTGTCGATCAGACAAAC 3360  
Qy 3361 GGTATGTCTTTTAAAGAGAGAGAGCTGGGCTCAATTTGCTGAGGAACTGCTTGGCGTGA 3420  
Db 3361 GGTATGTCTTTTAAAGAGAGAGAGCTGGGCTCAATTTGCTGAGGAACTGCTTGGCGTGA 3420  
Qy 3421 TTTGCTCCCTTCTGACCAATGTACAAAGAGAGAGTGGTTCGCGGCGGAGAGCTTACC 3480  
Db 3421 TTTGCTCCCTTCTGACCAATGTACAAAGAGAGAGTGGTTCGCGGCGGAGAGCTTACC 3480  
Qy 3481 CACATTAATGACTACCTGAATTTCTCATGCGCAGGCTGCGATGTAACAAGAACTGAA 3540  
Db 3481 CACATTAATGACTACCTGAATTTCTCATGCGCAGGCTGCGATGTAACAAGAACTGAA 3540  
Qy 3541 GCTTTCAACAATGCCATGAAGAGCGGCCAAGATATCAGAAAGCGGCTCATTTTGGAT 3600  
Db 3541 GCTTTCAACAATGCCATGAAGAGCGGCCAAGATATCAGAAAGCGGCTCATTTTGGAT 3600  
Qy 3601 CCGGATCTGCTTCTACTACAGTTCAGGTTCTTCAAGAGATTAAGCAAGTGGCGTCC 3660  
Db 3601 CCGGATCTGCTTCTACTACAGTTCAGGTTCTTCAAGAGATTAAGCAAGTGGCGTCC 3660  
Qy 3661 GCACTGCTATTCAGACTCTGAAGAAACCGTATCGCGAAGTCAAGAAATATGCGAG 3720  
Db 3661 GCACTGCTATTCAGACTCTGAAGAAACCGTATCGCGAAGTCAAGAAATATGCGAG 3720  
Qy 3721 TTGGTCAAAAACAAGAGATGAGAGACAGCAAGAACCCCTGCTCGCGTCAACAG 3780  
Db 3721 TTGGTCAAAAACAAGAGATGAGAGACAGCAAGAACCCCTGCTCGCGTCAACAG 3780  
Qy 3781 GTTTATGAATAAATGAGGCGCATCAGCGCTGAAGCGCATGAATAATCGGAGCAATTA 3840  
Db 3781 GTTTATGAATAAATGAGGCGCATCAGCGCTGAAGCGCATGAATAATCGGAGCAATTA 3840  
Qy 3841 GATTTCAAGGTGATCAGTTGCTGAGCTGTCTTCTCGCGACCGTGAATGAATACA 3900  
Db 3841 GATTTCAAGGTGATCAGTTGCTGAGCTGTCTTCTCGCGACCGTGAATGAATACA 3900  
Qy 3901 TGGGCAATGCTGAGGCTTACAGCGGCTTTCACACTGTACTACCAAGAGCCCAAGTTC 3960  
Db 3901 TGGGCAATGCTGAGGCTTACAGCGGCTTTCACACTGTACTACCAAGAGCCCAAGTTC 3960

QY 3961 GTGTGGCAGATGGCGGGCAGACAGCTCGCTACATTAAAGCCGACGATGACGACGACGCC 4020  
DB 3961 GTGTGGCAGATGGCGGGCAGACAGCTCGCTACATTAAAGCCGACGATGACGACGACGCC 4020  
QY 4021 GGTGAAGCGCCCGCGGCGTGTGATGACCGCGTTCATGATGACGCGGCTTGATGCGCGATTAAG 4080  
DB 4021 GGTGAAGCGCCCGCGGCGTGTGATGACCGCGTTCATGATGACGCGGCTTGATGCGCGATTAAG 4080  
QY 4081 AAGTTTACGAAGCAGTATGTGGCCAGGCTGGAAGGCGCATGATCGAGTACCTGTATCCG 4140  
DB 4081 AAGTTTACGAAGCAGTATGTGGCCAGGCTGGAAGGCGCATGATCGAGTACCTGTATCCG 4140  
QY 4141 GAGGTCTATGAAGTGTGGCGCGATGATTTGATGATGATTTGTTTCAAGGGAATGCG 4200  
DB 4141 GAGGTCTATGAAGTGTGGCGCGATGATTTGATGATGATTTGTTTCAAGGGAATGCG 4200  
QY 4201 GATTAT 4206  
DB 4201 GATTAT 4206

RESULT 2  
AX034425 8045 bp DNA linear PAT 22-SEP-2000  
LOCUS Sequence 1 from Patent WO0050581.  
DEFINITION AX034425  
ACCESSION AX034425.1 GI:10303130  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Neurospora crassa  
Neurospora crassa  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
1  
Cognoni, C. and Macino, G.  
Isolation and characterization of a n. Crassa silencing gene and  
uses thereof  
Patent: WO 0050581-A 1 31-AUG-2000;  
UNIV ROMA (IT) ; COGONI CARLO (IT) ; MACINO GIUSEPPE (IT)  
Location/Qualifiers  
1. 8045  
/organism="Neurospora crassa"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:5141"  
2447..6655  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAC10121.1"  
/translation="MNPITPRKRNSPVEEINRLNDVNLGLQVADTTLTPHREKL  
ABSDEDFGRDXTYRALNPLPYRKDSLNQAEANPTIEAKAASNNWPKAHADPTLP  
WSKEPPRAATAGQOMALQVLLEVLNKFMPNNPTGKRTGRTLSGSPSLSTNT  
KRDEPANTVPADPKRSLSRATGPRIHGAALPLKFPDVPNTGSKRPSLESNLNQC  
TKRAKGLSDNVAAAAAPVPIASALDKVTRRHANTDEPTAGHRAADQVDFDTSQ  
GTYSGVSVFASCHNOSTTOSPEAPSPORREKRPVDAVFEAGHLIESPSKRTKS  
HINOPSSSSOGESTSPSTYVESFPSSGSPAGAIPEPSASGLARSESASOVYAP  
VVAARLRNTYPRKPKMLHEAPLVAHEVTRPLFMHCYDLDESLGLKTPDSMTAADV  
TDLWKTLRLDAFRGKPFEPKPEPNDVFLVMTNPFESKSAVLSVLDPNPSPTA  
PLVLKTLKPLFEQGRCLTRRFGPDRFELLIPSPSTSPSVVSKPGAEVNIQ  
MLTMGHSILVGRMRAFFAKADGYRKLRFQRLAADPKFIIEKRVHAFETDITPRP  
DVFKTSVVAEPEVORCTEFKVSOMDLQDNNTWOHLTLFSRIQLGSKTYAI  
MTLEPHQIRHHTKDLSPSGTGVMNDGVRMSVAKRIRVTLGADVPSAOGSG  
SAGKMVIVIDVDTGDEDMITTPYSPQRKCEDPYDKORTLEVSVASELKSAGLNLQ  
LPLVEDBARDKVRKQAIIGRLINDLRQSEBOKHALNRFVERFROWYESYSRARV  
SHGRVPLLAGLPDSOEBTLNPLMNSGDFPKQRYLQDIANDLQKRCDTLKSRLNRY  
GSGAYLYMIDFWGLVEENHVHGFSGKFEDEESFTLSDCVLVARSAFHPDIO  
RLVAFKPEPLHSLKQVLIIPSTKGDVPLAKLSDGDDYDGMVACMPBEIYDGVNM  
PLEPDLRYLKDKTKTEPKOLMASHGSGAKKEQTTDMIOKSFHPLAOPMLGCMNY  
KEELCYINNSVSKRAIILISLIYANVDOSKQIVPRKASMAIDREBLGALSLEDP  
MYSDSMWLGSEPTTHIIDLYKFSIARPAIDIEKAFHNAMKADKEDQAPVAVDIA  
SYVTFPEKISDKSRSSALLEFTLKNRIGVEKEKYGRLVKNKEMRDKDQAPVAVDIA  
EKWAIITPEAMDKSGANYDSKVLRLLESLADREMTWALILASTAFKLZYHKSFKP  
VMQAGROLAYIKQMTSRPGEAPALMTAFMYAGLMPDKKFTKQYVARLEBGSSEYP  
DEVEYEVLDGDDDFDGIQFTNGDY"

ORIGIN  
Query Match 100.0%; Score 4206; DB 6; Length 8045;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4206; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAACCTTATTTACTCTTAGAAGAGAAATAGCCCCGTGAGGAATCTTAAACGGCTTC 60  
DB 2447 ATGAACCTTATTTACTCTTAGAAGAGAAATAGCCCCGTGAGGAATCTTAAACGGCTTC 2506  
QY 61 AATTAACGACTACAACTGGGCGCTCCAGTGTGTGACAGACAACTGACCCCCCGACCG 120  
DB 2507 AATTAACGACTACAACTGGGCGCTCCAGTGTGTGACAGACAACTGACCCCCCGACCG 2566  
QY 121 CGGAAGAGCTGGCCGAGAGTGAAGAGATTTTCGTGCGCATGACAAAGATCTTACAGAGCC 180  
DB 2567 CGGAAGAGCTGGCCGAGAGTGAAGAGATTTTCGTGCGCATGACAAAGATCTTACAGAGCC 2626  
QY 181 CTGAACCTTTCTCTACTGCGGGAAGATGACTCCCTTAACCGAGCGAAGCAAATTCTTC 240  
DB 2627 CTGAACCTTTCTCTACTGCGGGAAGATGACTCCCTTAACCGAGCGAAGCAAATTCTTC 2686  
QY 241 ATGAGGCCAAAGCTGCGAGCTCGAACTGGGTGCCCCAAAGCCGACGCGATGACAG 300  
DB 2687 ATGAGGCCAAAGCTGCGAGCTCGAACTGGGTGCCCCAAAGCCGACGCGATGACAG 2746  
QY 301 CTTTCGTGTGTCAGAGAACCTCCCGCGCGCTACTGCGGCGCAACATGGGCAATTGGAG 360  
DB 2747 CTTTCGTGTGTCAGAGAACCTCCCGCGCGCTACTGCGGCGCAACATGGGCAATTGGAG 2806  
QY 361 ACTGTCTTCTGAGAGTGTCTTAATAGTTTATGCGCACTTCCCAATPAACACAGGTGCA 420  
DB 2807 ACTGTCTTCTGAGAGTGTCTTAATAGTTTATGCGCACTTCCCAATPAACACAGGTGCA 2866  
QY 421 ACCTTTGGCAGAACTTAAAGCGGCCCAAGTGGCTGAGCGGCCCAACTCTACCAACACC 480  
DB 2867 ACCTTTGGCAGAACTTAAAGCGGCCCAAGTGGCTGAGCGGCCCAACTCTACCAACACC 2926  
QY 481 AAACGCAAGATGAGCGCCGCAATGTCACTTTTGCTGATTCGCGCCAAACGCTTGTTACT 540  
DB 2927 AAACGCAAGATGAGCGCCGCAATGTCACTTTTGCTGATTCGCGCCAAACGCTTGTTACT 2986  
QY 541 CGCTTCGCGCAGAGTCTCTTATTCAGCGCGGGCGCATACCCCTTAAGTTCCCGCATCCA 600  
DB 2987 CGCTTCGCGCAGAGTCTCTTATTCAGCGCGGGCGCATACCCCTTAAGTTCCCGCATCCA 3046  
QY 601 GGTGAATACCGGTTCACAAAGCAACATCTCTGAGAGTGAATCTCAATCAGTGACCAAG 660  
DB 3047 GGTGAATACCGGTTCACAAAGCAACATCTCTGAGAGTGAATCTCAATCAGTGACCAAG 3106  
QY 661 CGGCGCAAGGGCAAGCTGTGTGATTAATGTTGCGCGTGCAGCGCCGCGCTGCTATT 720  
DB 3107 CGGCGCAAGGGCAAGCTGTGTGATTAATGTTGCGCGTGCAGCGCCGCGCTGCTATT 3166  
QY 721 GCGAGGCGTTTGACAGAGTACCGACTCGAAGGCAATGCCAATPACAGAAATCCACAGGG 780  
DB 3167 GCGAGGCGTTTGACAGAGTACCGACTCGAAGGCAATGCCAATPACAGAAATCCACAGGG 3226  
QY 781 ACAGGTCAAGACGAGCGGACGAGTGAATCTCTTTGATATCTTCAAGGCACTTCTAT 840  
DB 3227 ACAGGTCAAGACGAGCGGACGAGTGAATCTCTTTGATATCTTCAAGGCACTTCTAT 3286  
QY 841 GGTTCGAGTGTCTTCAGCGCTTCCGCTCAATACAGAGCACTTACCAAGTAGATTGAG 900  
DB 3287 GGTTCGAGTGTCTTCAGCGCTTCCGCTCAATACAGAGCACTTACCAAGTAGATTGAG 9346  
QY 901 GCTCCCTCTTCAAGCGCCAGAGAAAGCGGCTGTGAGAGCCACGCTCTTACAGCTGGA 960  
DB 3347 GCTCCCTCTTCAAGCGCCAGAGAAAGCGGCTGTGAGAGCCACGCTCTTACAGCTGGA 3406  
QY 961 CACTTGAATTGAGTCTCTTACGAAAGGAAACAAACCAATGCCACATAGATTAACGAGCC 1020  
DB 3407 CACTTGAATTGAGTCTCTTACGAAAGGAAACAAACCAATGCCACATAGATTAACGAGCC 3466

QY 1021 CTTTCATGCTGCTTCCAGGGGTAACCTTGCTTCAAGCACTTACTATGAGTGGTTTCCAGT 1080  
DB 3467 CTTTCATGCTGCTTCCAGGGGTAACCTTGCTTCAAGCACTTACTATGAGTGGTTTCCAGT 3526  
QY 1081 TCCGCGCGCGAGGGCGCAATTCGCGAGCCGAGTGCCTCAATGGAAGTGGCTCGAGGCA 1140  
DB 3527 TCCGCGCGCGAGGGCGCAATTCGCGAGCCGAGTGCCTCAATGGAAGTGGCTCGAGGCA 3586  
QY 1141 GAAAGCGCTCGATCTCAAGTTCAAGTTCAATGCTCCCGGTGGTTGCAAGTCCGCTGAAGAAAT 1200  
DB 3587 GAAAGCGCTCGATCTCAAGTTCAAGTTCAATGCTCCCGGTGGTTGCAAGTCCGCTGAAGAAAT 3646  
QY 1201 ATTTGGCGGAAATTTCCCAATGGCTACAGAACTCCTCGCTGGTTGCAATGGGAAT 1260  
DB 3647 ATTTGGCGGAAATTTCCCAATGGCTACAGAACTCCTCGCTGGTTGCAATGGGAAT 3706  
QY 1261 ACCAGACTCTTTATGCACTGCAAGTAGACTTGGAAAGCAAGAGCTTGAGCTTAAAGTAC 1320  
DB 3707 ACCAGACTCTTTATGCACTGCAAGTAGACTTGGAAAGCAAGAGCTTGAGCTTAAAGTAC 3766  
QY 1321 GACCTTCTGCTGCTACCGCGCGAGTGTCAAGATATCTGAAAGACTCTTACCGGCTT 1380  
DB 3767 GACCTTCTGCTGCTACCGCGCGAGTGTCAAGATATCTGAAAGACTCTTACCGGCTT 3826  
QY 1381 GATGCTTTCCGCTGTAACCCCTTCCAGAAAGCGCCCAAGAGCTGTTGTAAGCGCA 1440  
DB 3827 GATGCTTTCCGCTGTAACCCCTTCCAGAAAGCGCCCAAGAGCTGTTGTAAGCGCA 3886  
QY 1441 ATGACGGGCAACTTTGAGAGCAAGTAGTGCCTGTTCTCTGCTGTTCTAGACTAC 1500  
DB 3887 ATGACGGGCAACTTTGAGAGCAAGTAGTGCCTGTTCTCTGCTGTTCTAGACTAC 3946  
QY 1501 AATCCGGAACAATCGGCTTACTGCGCCCTTTACTTGTGAAGCTGAAAGCGCTCATGTTTC 1560  
DB 3947 AATCCGGAACAATCGGCTTACTGCGCCCTTTACTTGTGAAGCTGAAAGCGCTCATGTTTC 4006  
QY 1561 GAGAGGGCTGTGCACTCAACCGTGGTGGTGGTCTGTAAGTGGTTTTCGAGATCCTTATA 1620  
DB 4007 GAGAGGGCTGTGCACTCAACCGTGGTGGTGGTCTGTAAGTGGTTTTCGAGATCCTTATA 4066  
QY 1621 CCGTGCCTTACGAGCAACGAGCCCAAGTGTACCGCGGTGTGACGAAACCAACGAGTGC 1680  
DB 4067 CCGTGCCTTACGAGCAACGAGCCCAAGTGTACCGCGGTGTGACGAAACCAACGAGTGC 4126  
QY 1681 GTCGAAGAATCATCTCAAGTGTGTCAAGTGGGCAACTTCTTTGTAAGCCGCCAATGG 1740  
DB 4127 GTCGAAGAATCATCTCAAGTGTGTCAAGTGGGCAACTTCTTTGTAAGCCGCCAATGG 4186  
QY 1741 CGGCTTTCTTCCGCAAGATGCGGGAATCAAGAAACCTCTCAAGGAGTTTCAGTCCGC 1800  
DB 4187 CGGCTTTCTTCCGCAAGATGCGGGAATCAAGAAACCTCTCAAGGAGTTTCAGTCCGC 4246  
QY 1801 GCCGAGGACCCGAAACCATCATCAAGAGAGAGTGCACCTTCTTTGCGGAGCCGCAAT 1860  
DB 4247 GCCGAGGACCCGAAACCATCATCAAGAGAGAGTGCACCTTCTTTGCGGAGCCGCAAT 4306  
QY 1861 AGCTTCCGACTGATGTGTTCAAGACGAGATCTGCTTCCGCGAGAGAACTGTAGAG 1920  
DB 4307 AGCTTCCGACTGATGTGTTCAAGACGAGATCTGCTTCCGCGAGAGAACTGTAGAG 4366  
QY 1921 CAAGGAGCCGAGTTCAAGTTAGTCAATGCTGAGCTGGCTCTGCAACCTGCAACAAC 1980  
DB 4367 CAAGGAGCCGAGTTCAAGTTAGTCAATGCTGAGCTGGCTCTGCAACCTGCAACAAC 4426  
QY 1981 ACTTGGCAGCCGCACTCTCAAGTTGTTCTCCGTAATCAAGCTCGGTCTGAGTAAGCATAT 2040  
DB 4427 ACTTGGCAGCCGCACTCTCAAGTTGTTCTCCGTAATCAAGCTCGGTCTGAGTAAGCATAT 4486  
QY 2041 GCCATTATGACATTGAGGCTGTCAAGATCAAGACCAAGACCGATCTTCTTCAACT 2100  
DB 4487 GCCATTATGACATTGAGGCTGTCAAGATCAAGACCAAGACCGATCTTCTTCAACT 4546

QY 2101 TCAGGCACTGGCGAAGTATGATGACGAGTGTAGGCGGCAATGTGCGCAAGCGTGCCAAAG 2160  
DB 4547 TCAGGCACTGGCGAAGTATGATGACGAGTGTAGGCGGCAATGTGCGCAAGCGTGCCAAAG 4606  
QY 2161 AGGATACGCGATGTTCTCGGTTTGGGTGATGTGCCCTCTGCTGTGCAAGGCGGTTTGGT 2220  
DB 4607 AGGATACGCGATGTTCTCGGTTTGGGTGATGTGCCCTCTGCTGTGCAAGGCGGTTTGGT 4666  
QY 2221 TCCGCGCAAGGGAATGTCGGGTTATCGACGTTGACGACACAGGCGATGAGATTGGATGAG 2280  
DB 4667 TCCGCGCAAGGGAATGTCGGGTTATCGACGTTGACGACACAGGCGATGAGATTGGATGAG 4726  
QY 2281 ACATACCCGTCGCGAGCGCAAGTGGAAATGCGACTTCTGTTGATAACATCAACGTAACCTC 2340  
DB 4727 ACATACCCGTCGCGAGCGCAAGTGGAAATGCGACTTCTGTTGATAACATCAACGTAACCTC 4786  
QY 2341 GAAATCCGAGCGTGCTTCTGAACTGAAATCACTGAGTCAAGCTGTTAATCCT 2400  
DB 4787 GAAATCCGAGCGTGCTTCTGAACTGAAATCACTGAGTCAAGCTGTTAATCCT 4846  
QY 2401 GTCCTGGAATATGAGCCAGGAGCAAGTGAAGATGGCCAGGCAATCGGTCACCGCTT 2460  
DB 4847 GTCCTGGAATATGAGCCAGGAGCAAGTGAAGATGGCCAGGCAATCGGTCACCGCTT 4906  
QY 2461 ATCAACGATTTTGCACACGACGTTTACGCGCAAAAGCATCTTGAATCCGCCAGTGGAA 2520  
DB 4907 ATCAACGATTTTGCACACGACGTTTACGCGCAAAAGCATCTTGAATCCGCCAGTGGAA 4966  
QY 2521 TTTGCGCAATGGGTTTACGAGGTTATTCAGTGCAGCACTGAGTCAAGCCAGCCCT 2580  
DB 4967 TTTGCGCAATGGGTTTACGAGGTTATTCAGTGCAGCACTGAGTCAAGCCAGCCCT 5026  
QY 2581 GTGCTTTTCTGCTGGGCTTACCTGACAGTCAAGAGAGACATGAACTTCTGATGAAAC 2640  
DB 5027 GTGCTTTTCTGCTGGGCTTACCTGACAGTCAAGAGAGACATGAACTTCTGATGAAAC 5086  
QY 2641 AGTGGGTTGATCCCAAGAACAAAGTACTTGCACACATTCGCTGGATCTTCAAAAG 2700  
DB 5087 AGTGGGTTGATCCCAAGAACAAAGTACTTGCACACATTCGCTGGATCTTCAAAAG 5146  
QY 2701 CGGAAATGTGACAGTGTGAAGTCCAGCTGAACATCCGTGTGATGATGATCAATTA 2760  
DB 5147 CGGAAATGTGACAGTGTGAAGTCCAGCTGAACATCCGTGTGATGATGATCAATTA 5206  
QY 2761 TACATGATTCGCAATTTCTGCGGTGCTGAGGAAATGAGGTTATGTCGATTCGC 2820  
DB 5207 TACATGATTCGCAATTTCTGCGGTGCTGAGGAAATGAGGTTATGTCGATTCGC 5266  
QY 2821 TCAAGTTTCAGAGACGAGAGAGTCTTTTACACTCTTATCGACTGTGATGTCTCGTG 2880  
DB 5267 TCAAGTTTCAGAGACGAGAGAGTCTTTTACACTCTTATCGACTGTGATGTCTCGTG 5326  
QY 2881 GCGGATTCGCCAGCCATTTTCCCTAGTATATCCAGCGGTTGAGCATCTTCAAGCCA 2940  
DB 5327 GCGGATTCGCCAGCCATTTTCCCTAGTATATCCAGCGGTTGAGCATCTTCAAGCCA 5386  
QY 2941 GAGCTTCACAGTCTCAAGATGTATAATCTTCTCTATCAAAAGGAGTATACGCTTGGT 3000  
DB 5387 GAGCTTCACAGTCTCAAGATGTATAATCTTCTCTATCAAAAGGAGTATACGCTTGGT 5446  
QY 3001 AAGAAGCTATCTGGTGTGAGACTACGAGCGGATATGCGCTGTGCTGTGGATCCGAG 3060  
DB 5447 AAGAAGCTATCTGGTGTGAGACTACGAGCGGATATGCGCTGTGCTGTGGATCCGAG 5506  
QY 3061 ATGCTGATGGTTTGTGCAATGCGAAATGCTCTGAGGCCGACCTGTCTAGGTACTTA 3120  
DB 5507 ATGCTGATGGTTTGTGCAATGCGAAATGCTCTGAGGCCGACCTGTCTAGGTACTTA 5566  
QY 3121 AAGAAGGACAAAGCACTTTTCAACCACTTATGGGCTCAACGCGCACGGGCTCAGCGGCC 3180  
DB 5567 AAGAAGGACAAAGCACTTTTCAACCACTTATGGGCTCAACGCGCACGGGCTCAGCGGCC 5626  
QY 3181 AAGAGCAGACTACATACGATATGATTCAGAAAGACTTTCATTTTGGCTTGCACGCCAAC 3240

Db 5627 AAGAGAGAGCTATCATGATATATCCAGAAAGCTTCCATTTCCGCTCCAGCCCAAC 5686  
Qy 3241 TTCTTGGGAGATGTCACCTAACTAACAAAGAAAGGCTGTTAACTCAATCAATAGTGTCT 3300  
Db 5687 TTCTTGGGAGATGTCACCTAACTAACAAAGAAAGGCTGTTAACTCAATAGTGTCT 5746  
Qy 3301 AACAAAGCCGGCCATCATTTCTTAACTGTCAGTGTGGGAAACCTCGTCAATCAGAGCAAC 3360  
Db 5747 AACAAAGCCGGCCATCATTTCTTAACTGTCAGTGTGGGAAACCTCGTCAATCAGAGCAAC 5806  
Qy 3361 GGTATTTCTCTTAACTGTCAGTGTGGGCTCAATGGCTAGGAACTGTGGGGTGTCA 3420  
Db 5807 GGTATTTCTCTTAACTGTCAGTGTGGGCTCAATGGCTAGGAACTGTGGGGTGTCA 5866  
Qy 3421 TTGCTCCCTTCTGACCAATGTCAGAGGCAAGTGGCTCGGGCCGGAGAGCTTACC 3480  
Db 5867 TTGCTCCCTTCTGACCAATGTCAGAGGCAAGTGGCTCGGGCCGGAGAGCTTACC 5926  
Qy 3481 CACATTTATGACTACCTGAAATTTCTCATGCGCAGGCTCGATTTGACAAAGAACTGGAA 3540  
Db 5927 CACATTTATGACTACCTGAAATTTCTCATGCGCAGGCTCGATTTGACAAAGAACTGGAA 5986  
Qy 3541 GCCTTCCAAATGTCAGTGTGGGCTCAATGGCTAGGAACTGTGGGGTGTCA 3600  
Db 5987 GCCTTCCAAATGTCAGTGTGGGCTCAATGGCTAGGAACTGTGGGGTGTCA 6046  
Qy 3601 CCGGATCTCGCTTCTCTACTACAGCTTCTTCAAGAGATTTAGCAGCAAGTGCAGTCC 3660  
Db 6047 CCGGATCTCGCTTCTCTACTACAGCTTCTTCAAGAGATTTAGCAGCAAGTGCAGTCC 6106  
Qy 3661 GCACCTGCTATTCACGACTCTGAGAACCGTATCGGCGAAGTCGAGAAAGATATGCGAGG 3720  
Db 6107 GCACCTGCTATTCACGACTCTGAGAACCGTATCGGCGAAGTCGAGAAAGATATGCGAGG 6166  
Qy 3721 TTGGTCAAAAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3780  
Db 6167 TTGGTCAAAAGAGAGATGAG 6226  
Qy 3781 GTTTATGAAAAATGTCGCGCATCAGCCTGAGGCGATGAGCAAAATCCGAGCAAAATTTAT 3840  
Db 6227 GTTTATGAAAAATGTCGCGCATCAGCCTGAGGCGATGAGCAAAATCCGAGCAAAATTTAT 6286  
Qy 3841 GATTCTAAGGTATCAGGTGCTGAGAGCTGCTCTTCTCGCGAGCCGTGAGATGATATCA 3900  
Db 6287 GATTCTAAGGTATCAGGTGCTGAGAGCTGCTCTTCTCGCGAGCCGTGAGATGATATCA 6346  
Qy 3901 TTGGGCAATTTGCTGAGGAGCTAGAGGCTTTCAAGCTGTAACAACAAGAGAGAGAGAGAG 3960  
Db 6347 TTGGGCAATTTGCTGAGGAGCTAGAGGCTTTCAAGCTGTAACAACAAGAGAGAGAGAGAG 6406  
Qy 3961 GTGTGCGAGATGCGGCGAGAGAGCTCGCTACCTTAAGGCGAGATGAGAGAGAGAGAG 4020  
Db 6407 GTGTGCGAGATGCGGCGAGAGAGCTCGCTACCTTAAGGCGAGATGAGAGAGAGAGAGAG 6466  
Qy 4021 GGTGAAGAGGCGCGCGGCTGTTGATGACCGGCTTCAATGATGCGGCTTGTGATGCGGATTAAG 4080  
Db 6467 GGTGAAGAGGCGCGCGGCTGTTGATGACCGGCTTCAATGATGCGGCTTGTGATGCGGATTAAG 6526  
Qy 4081 AAGTTTACGAAGAGATGATGTCAGGCTGAGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAG 4140  
Db 6527 AAGTTTACGAAGAGATGATGTCAGGCTGAGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 6586  
Qy 4141 GAGGCTATGAGAGAGTGTGGGAGATGATTTTGTGAGAAATGTTTCAAGAGAGAGAGAGAG 4200  
Db 6587 GAGGCTATGAGAGAGTGTGGGAGATGATTTTGTGAGAAATGTTTCAAGAGAGAGAGAGAG 6646  
Qy 4201 GATTAT 4206  
Db 6647 GATTAT 6652

RESULT 3

CO803434 4206 bp DNA linear PAT 10-MAY-2004  
LOCUS CO803434  
DEFINITION Sequence 1 from Patent WO2004035784.  
ACCESSION CO803434  
VERSION CO803434.1 GI:47110295  
KEYWORDS  
SOURCE Neurospora crassa  
ORGANISM Neurospora crassa  
REFERENCE 1  
AUTHORS Mekeyev, E. and Bamford, D.  
TITLE Soluble rna polymerase protein and methods for the use thereof  
JOURNAL Patent: WO 2004035784-A 1 29-APR-2004;  
RNA-line Oy (FI)  
FEATURES  
source 1..4206  
Location/Qualifiers  
/organism="Neurospora crassa"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:5141"  
ORIGIN  
Query Match 99.9%; Score 4202.8; DB 6; Length 4206;  
Best Local Similarity 100.0%; Pred. No. 0; Mismatches 2; Indels 0; Gaps 0;  
Matches 4204; Conservative 0;  
Qy 1 ATGAACTTATTAATCTCTAGAGAGAGAAATAGCCCGTCGAGGAATCATTAACCGGCTC 60  
Db 1 ATGAACTTATTAATCTCTAGAGAGAGAAATAGCCCGTCGAGGAATCATTAACCGGCTC 60  
Qy 61 AATTAAGACTTAACTAGGAGCTTCCAGTGTGTGACAGACAACTTACCCCCACCGC 120  
Db 61 AATTAAGACTTAACTAGGAGCTTCCAGTGTGTGACAGACAACTTACCCCCACCGC 120  
Qy 121 CGGAGAGAGCTGGCGAGAGTGAAGAGATTTGCGTGGCATAGCAAGATCTTACAGAGCC 180  
Db 121 CGGAGAGAGCTGGCGAGAGTGAAGAGATTTGCGTGGCATAGCAAGATCTTACAGAGCC 180  
Qy 181 CTGAACCTTCTCTACTGCGGAGAGATGACTCCCTGTAACAGAGCAGAGCAACTTCTTC 240  
Db 181 CTGAACCTTCTCTACTGCGGAGAGATGACTCCCTGTAACAGAGCAGAGCAACTTCTTC 240  
Qy 241 ATGAGAGCCAAAGCTGCGAGCTTGAATCTGGTGTGCCAAAGCCGACCTTGACAGC 300  
Db 241 ATGAGAGCCAAAGCTGCGAGCTTGAATCTGGTGTGCCAAAGCCGACCTTGACAGC 300  
Qy 301 CTTCCGTTGTCGAAGGAACTCCCGGCGGCTGACTGCGGCGCAACATGGGCAATTGCGAG 360  
Db 301 CTTCCGTTGTCGAAGGAACTCCCGGCGGCTGACTGCGGCGCAACATGGGCAATTGCGAG 360  
Qy 361 ACTGTGTTGCTCGAGGTGCTTAATAGTTTATGCCACTCCCAATTAACACCAAGGTGCA 420  
Db 361 ACTGTGTTGCTCGAGGTGCTTAATAGTTTATGCCACTCCCAATTAACACCAAGGTGCA 420  
Qy 421 AGCTTTGGCAGAACTTAAGCGGCGCAAGTGGCTGAGCGGCCCAACTCTTACCAACACC 480  
Db 421 AGCTTTGGCAGAACTTAAGCGGCGCAAGTGGCTGAGCGGCCCAACTCTTACCAACACC 480  
Qy 481 AAAAGCAAGGATGAGCGCGCAATGTCATTTGCTGATGCGGCGCAACGCTGTTGACT 540  
Db 481 AAAAGCAAGGATGAGCGCGCAATGTCATTTGCTGATGCGGCGCAACGCTGTTGACT 540  
Qy 541 CGCTTTCGCAAGAGTCTCTATTTCAGGCGCGGCGAGTACCCCTTAAGTTTCCCGATCCA 600  
Db 541 CGCTTTCGCAAGAGTCTCTATTTCAGGCGCGGCGAGTACCCCTTAAGTTTCCCGATCCA 600  
Qy 601 GTGAATACCGGTTCCAAAGCAATCTTTCGAGAGTGAAGATTCATATGATGACCAAG 660  
Db 601 GTGAATACCGGTTCCAAAGCAATCTTTCGAGAGTGAAGATTCATATGATGACCAAG 660  
Qy 661 CGGCGCAAGGCGAAGGCTGATTAATGTTGCGGCTGCGGCGCGCGCGCGCGGCTATT 720  
Db 661 CGGCGCAAGGCGAAGGCTGATTAATGTTGCGGCTGCGGCGCGCGCGCGCGGCTATT 720

QY 721 GCGAGCGCTTTGACAGAGTACCGAGCTCGAAGGCGCAATACAGAGATCCACGCGG 780  
DB 721 GCGAGCGCTTTGACAGAGTACCGAGCTCGAAGGCGCAATACAGAGATCCACGCGG 780  
QY 781 ACAGGTCATAGACGAGCGGACCGAGTGGATTCCTTTGATACATCTCAAGGCACTTCTAT 840  
DB 781 ACAGGTCATAGACGAGCGGACCGAGTGGATTCCTTTGATACATCTCAAGGCACTTCTAT 840  
QY 841 GGTTCGAGTGTCTTACGCGCTTCCGCTCAATACAGAGTACCGAGAGTGGATTTGAG 900  
DB 841 GGTTCGAGTGTCTTACGCGCTTCCGCTCAATACAGAGTACCGAGAGTGGATTTGAG 900  
QY 901 GGTTCGAGTGTCTTACGCGCTTCCGCTCAATACAGAGTACCGAGAGTGGATTTGAG 960  
DB 901 GGTTCGAGTGTCTTACGCGCTTCCGCTCAATACAGAGTACCGAGAGTGGATTTGAG 960  
QY 961 CACTTGATAGTCTCTAGCAAGAGAAACAACCAAGTCCACATAGATACCAAGCC 1020  
DB 961 CACTTGATAGTCTCTAGCAAGAGAAACAACCAAGTCCACATAGATACCAAGCC 1020  
QY 1021 CTTTCATGCTCTTCCAGGAGTAACTTCTTACAGACTTATATAGTGGTTTCAAGT 1080  
DB 1021 CTTTCATGCTCTTCCAGGAGTAACTTCTTACAGACTTATATAGTGGTTTCAAGT 1080  
QY 1081 TCCGCGCGGCGAGGCGGCAATCCCGAGCGGAGTGGCTCAATAGGCTGCGGAGGAA 1140  
DB 1081 TCCGCGCGGCGAGGCGGCAATCCCGAGCGGAGTGGCTCAATAGGCTGCGGAGGAA 1140  
QY 1141 GAAAGCGCTGATCTCAAGTTCATGCTCCCGGAGTGGAGTCCGCTGAGAAAT 1200  
DB 1141 GAAAGCGCTGATCTCAAGTTCATGCTCCCGGAGTGGAGTCCGCTGAGAAAT 1200  
QY 1201 ATTTGGCCGAATTTCCCAATGGCTTACAGAGTCTCTCGCTGTTGCAATGGAAAT 1260  
DB 1201 ATTTGGCCGAATTTCCCAATGGCTTACAGAGTCTCTCGCTGTTGCAATGGAAAT 1260  
QY 1261 ACCGAGCTTTTATGCACTGCAAGTATGAGTGGAGCGAGGCTGGGCTTAAAGTAC 1320  
DB 1261 ACCGAGCTTTTATGCACTGCAAGTATGAGTGGAGCGAGGCTGGGCTTAAAGTAC 1320  
QY 1321 GACCTCTCTGCTTACCGCGCGGAGTGTCAAGATATCTGGAAGACTCTTACCGGCT 1380  
DB 1321 GACCTCTCTGCTTACCGCGCGGAGTGTCAAGATATCTGGAAGACTCTTACCGGCT 1380  
QY 1381 GATGCTTTCCGTGGTAAACCTTTCCAGAAAAGCGCCCAAGAGTGTGTGACCGCA 1440  
DB 1381 GATGCTTTCCGTGGTAAACCTTTCCAGAAAAGCGCCCAAGAGTGTGTGACCGCA 1440  
QY 1441 ATGACGGGCAACTTTGAGAGCAAGTATGAGCGGTGTTCTGCTGCTTCTAGACTAC 1500  
DB 1441 ATGACGGGCAACTTTGAGAGCAAGTATGAGCGGTGTTCTGCTGCTTCTAGACTAC 1500  
QY 1501 AATCCGAGCAACTGCGCTTACGCGCCCTTACCTTGTGAGTGAAGCGCTCATGTTC 1560  
DB 1501 AATCCGAGCAACTGCGCTTACGCGCCCTTACCTTGTGAGTGAAGCGCTCATGTTC 1560  
QY 1561 GAGCAGGAGCTGTGCACTCAACCGTGGTTCGCTCTGATAGTGGTTTTCAGATCTTATA 1620  
DB 1561 GAGCAGGAGCTGTGCACTCAACCGTGGTTCGCTCTGATAGTGGTTTTCAGATCTTATA 1620  
QY 1621 CCGTGGCTTACGAGCAAGCGCAAGTGTACCGCGGCTGTCAGCAAAACAACGAGTGG 1680  
DB 1621 CCGTGGCTTACGAGCAAGCGCAAGTGTACCGCGGCTGTCAGCAAAACAACGAGTGG 1680  
QY 1681 GTCAAGAGTCACTCAAGTGGCTCAAGTGGGCAACTTCTTGTATAGCGCGCAATGG 1740  
DB 1681 GTCAAGAGTCACTCAAGTGGCTCAAGTGGGCAACTTCTTGTATAGCGCGCAATGG 1740  
QY 1741 GCGCGCTTTCTTGGCAAGATGCGGATACAGAAAACCTTCAAGGAGTTCACGCTCGC 1800  
DB 1741 GCGCGCTTTCTTGGCAAGATGCGGATACAGAAAACCTTCAAGGAGTTCACGCTCGC 1800

QY 1801 GCGAGAGCCCGAAACCATCATCAAGAGAGTCCACTTCTTTCGCGAGACCGGACT 1860  
DB 1801 GCGAGAGCCCGAAACCATCATCAAGAGAGTCCACTTCTTTCGCGAGACCGGACT 1860  
QY 1861 AGCTTCGAGCTGATGTGTTCAAGAGAGTCTGTGTTCCGCAAGAGAACTGTAGAG 1920  
DB 1861 AGCTTCGAGCTGATGTGTTCAAGAGAGTCTGTGTTCCGCAAGAGAACTGTAGAG 1920  
QY 1921 CAACGAGCGAGTTCAGAGTATCAAAATGCTGAGACTGCTCTGCAACTCGAACAC 1980  
DB 1921 CAACGAGCGAGTTCAGAGTATCAAAATGCTGAGACTGCTCTGCAACTCGAACAC 1980  
QY 1981 ACTTGGCAGCGGACCTCAAGTGTCTCCGATACAGCTCGGCTGATAGACATAT 2040  
DB 1981 ACTTGGCAGCGGACCTCAAGTGTCTCCGATACAGCTCGGCTGATAGACATAT 2040  
QY 2041 GCCATTTATGACATTTGAGGCTTACACAGATCAGACACCAAGACCGATCTTTCACCT 2100  
DB 2041 GCCATTTATGACATTTGAGGCTTACACAGATCAGACACCAAGACCGATCTTTCACCT 2100  
QY 2101 TCAGGCTGCGGAGAGTGAATGACGCTGATAGCGCGGATGTGCGAAGCTGCGGCAAG 2160  
DB 2101 TCAGGCTGCGGAGAGTGAATGACGCTGATAGCGCGGATGTGCGAAGCTGCGGCAAG 2160  
QY 2161 AGGATACGAGTGTCTCGGCTTGGGATGTGCTCTGCTGCAAGGCGGTTGGT 2220  
DB 2161 AGGATACGAGTGTCTCGGCTTGGGATGTGCTCTGCTGCAAGGCGGTTGGT 2220  
QY 2221 TCGGCGAAGGAAATGTTGGTTATGACGTTGACGACACAGGCGATGAGATGGATCGAG 2280  
DB 2221 TCGGCGAAGGAAATGTTGGTTATGACGTTGACGACACAGGCGATGAGATGGATCGAG 2280  
QY 2281 ACATACCCGCTCCAGCGCAAGTGGAAATGCGACTTGTATTAACATCAACGTAACCTC 2340  
DB 2281 ACATACCCGCTCCAGCGCAAGTGGAAATGCGACTTGTATTAACATCAACGTAACCTC 2340  
QY 2341 GAAATCCGAGCGTGTCTGAACTGAACTGAGTCTCAACTACAGCTGTTACT 2400  
DB 2341 GAAATCCGAGCGTGTCTGAACTGAACTGAGTCTCAACTACAGCTGTTACT 2400  
QY 2401 GTCTGAGAGATGAGGCAAGGCAAGTGAAGTGGCCAGCAATCGGTGACCGTCT 2460  
DB 2401 GTCTGAGAGATGAGGCAAGGCAAGTGAAGTGGCCAGCAATCGGTGACCGTCT 2460  
QY 2461 ATCAAGATTTGAAACGACGTTCAAGGCAAGAAAGTCTTGAATGCGCCAGTGGAA 2520  
DB 2461 ATCAAGATTTGAAACGACGTTCAAGGCAAGAAAGTCTTGAATGCGCCAGTGGAA 2520  
QY 2521 TTTGCGCAATGGGTTTACGAGATTAATTCAGTGCAGCACTGAGTCAAGCGCGT 2580  
DB 2521 TTTGCGCAATGGGTTTACGAGATTAATTCAGTGCAGCACTGAGTCAAGCGCGT 2580  
QY 2581 GTGCTTTTCTGCTGGCTTACCTGACGTTCAAGAGGACACTGAACTTCTTATGAGAC 2640  
DB 2581 GTGCTTTTCTGCTGGCTTACCTGACGTTCAAGAGGACACTGAACTTCTTATGAGAC 2640  
QY 2641 AGTGGTTCATCCCAAGAGCAAAAGTACTTGCAGACATTCGCTGGGATCTTCAAAAG 2700  
DB 2641 AGTGGTTCATCCCAAGAGCAAAAGTACTTGCAGACATTCGCTGGGATCTTCAAAAG 2700  
QY 2701 CGGAAATGTGACAGTGAAGTCAAGCTGAACATCGGTGCTGATAGCATATATT 2760  
DB 2701 CGGAAATGTGACAGTGAAGTCAAGCTGAACATCGGTGCTGATAGCATATATT 2760  
QY 2761 TACATGATTCGCAATTTCTGGGCTGCTTACAGGAAATAGGTTCTGCGATTTCTCC 2820  
DB 2761 TACATGATTCGCAATTTCTGGGCTGCTTACAGGAAATAGGTTCTGCGATTTCTCC 2820  
QY 2821 TCAAGTTCAGACAGGAGGAGTCTTTTACACTCTTATCGGACTGTGATGCTCTG 2880  
DB 2821 TCAAGTTCAGACAGGAGGAGTCTTTTACACTCTTATCGGACTGTGATGCTCTG 2880  
QY 2881 GCGGATCCCAAGCCCATTTCTAGTATCAACGAGGTGAGAGTCTTCAAGCA 2940

Db 2881 GCGGAGATCCCGAGCCATTTCCATGATATCCAGGGGTCGAGAGCTTCAAGCCA 2940  
Qy 2941 GAGCTCAGAGCTCAAGAGATGATATCATCTTCTCTAAGAGATGATACGGCTTGT 3000  
Db 2941 GAGCTCAGAGCTCAAGAGATGATATCATCTTCTCTAAGAGATGATACGGCTTGT 3000  
Qy 3001 AAGAGATCTAGTGGAGAGATCAAGAGAGATATGGAGTGGGTCGAGTGGAG 3060  
Db 3001 AAGAGATCTAGTGGAGAGATCAAGAGAGATATGGAGTGGGTCGAGTGGAG 3060  
Qy 3061 ATGTCATGATGTTTCTGATATGGGAAATGCTCTGAGAGCCGACCTGTAGTACTTA 3120  
Db 3061 ATGTCATGATGTTTCTGATATGGGAAATGCTCTGAGAGCCGACCTGTAGTACTTA 3120  
Qy 3121 AAGAGAGCAAAAGACTTTCAAGCAACTATGAGCCCTCAGAGGAGAGGAGGAGCC 3180  
Db 3121 AAGAGAGCAAAAGACTTTCAAGCAACTATGAGCCCTCAGAGGAGAGGAGGAGCC 3180  
Qy 3181 AAGAGAGCACTACATATGATATGATCAGAGAGCTTCCATTTCCGCTCAGAGCCCAAC 3240  
Db 3181 AAGAGAGCACTACATATGATATGATCAGAGAGCTTCCATTTCCGCTCAGAGCCCAAC 3240  
Qy 3241 TTCTTGGGAGATGAGCACTATCAAGAGAGAGGCTCTGTATCATCAACATATGTTCT 3300  
Db 3241 TTCTTGGGAGATGAGCACTATCAAGAGAGAGGCTCTGTATCATCAACATATGTTCT 3300  
Qy 3301 AAGAGAGGAGCACTATCTTCTAGTCTAGTGGAGAAACCTCTGATCAGAGCAAGCAA 3360  
Db 3301 AAGAGAGGAGCACTATCTTCTAGTCTAGTGGAGAAACCTCTGATCAGAGCAAGCAA 3360  
Qy 3361 GGTATGTCCTTTAAGAGAGAGAGGAGGCTCAATGGTGGAGAACTGTTGGCGGTGA 3420  
Db 3361 GGTATGTCCTTTAAGAGAGAGAGGAGGCTCAATGGTGGAGAACTGTTGGCGGTGA 3420  
Qy 3421 TTGTCCCTTCTAAGCAAGATGTAAGAGAGAGAGTGGTCCGGGCGGAGAGCCCTACC 3480  
Db 3421 TTGTCCCTTCTAAGCAAGATGTAAGAGAGAGAGTGGTCCGGGCGGAGAGCCCTACC 3480  
Qy 3481 CACATTAATGATACCTGAATTTCTCCATCGCAGAGCTCGATTCAGAGAACTGGA 3540  
Db 3481 CACATTAATGATACCTGAATTTCTCCATCGCAGAGCTCGATTCAGAGAACTGGA 3540  
Qy 3541 GCCTTCACAAATGCCATGAAGAGGCGCAAGATACAGAGAGCGCGCTCACTTTGGAT 3600  
Db 3541 GCCTTCACAAATGCCATGAAGAGGCGCAAGATACAGAGAGCGCGCTCACTTTGGAT 3600  
Qy 3601 CCGGATTCGCTTCTAATCAGCTTCTTCAAGAGATTAAGCAAGTGGCATGCTCC 3660  
Db 3601 CCGGATTCGCTTCTAATCAGCTTCTTCAAGAGATTAAGCAAGTGGCATGCTCC 3660  
Qy 3661 GCATGCTATTCAGACTCTGAAGAACCTGATCGGAGAACTCGAAGAAATATGCGAG 3720  
Db 3661 GCATGCTATTCAGACTCTGAAGAACCTGATCGGAGAACTCGAAGAAATATGCGAG 3720  
Qy 3721 TTGCTCAAAACAGAGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3780  
Db 3721 TTGCTCAAAACAGAGAGATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3780  
Qy 3781 GTTTATGAAGAAATGAGGAGATCAGAGCTTGAAGAGAGAGAGAGAGAGAGAGAG 3840  
Db 3781 GTTTATGAAGAAATGAGGAGATCAGAGCTTGAAGAGAGAGAGAGAGAGAGAGAG 3840  
Qy 3841 GATTTAAGGTGATCAGAGTCTGAGAGCTTCTTCTCGAGAGAGAGAGAGAGAGAG 3900  
Db 3841 GATTTAAGGTGATCAGAGTCTGAGAGCTTCTTCTCGAGAGAGAGAGAGAGAGAG 3900  
Qy 3901 TGGGCAATGCTGAGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3960  
Db 3901 TGGGCAATGCTGAGAGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3960  
Qy 3961 GTTGTGAGATGAGGAG 4020  
Db 3961 GTTGTGAGATGAGGAG 4020

Db 3961 GTTGTGAGATGAGGAG 4020  
Qy 4021 GGTGAAG 4080  
Db 4021 GGTGAAG 4080  
Qy 4081 AAGTTACAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4140  
Db 4081 AAGTTACAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 4140  
Qy 4141 GAGGCTATGAAGATGAG 4200  
Db 4141 GAGGCTATGAAGATGAG 4200  
Qy 4201 GATTAT 4206  
Db 4201 GATTAT 4206

RESULT 4  
CO803436 3078 bp DNA linear PAT 10-MAY-2004  
LOCUS  
DEFINITION  
Sequence 3 from Patent WO2004035784.  
ACCESSION  
CO803436  
VERSION  
CO803436.1 GI:47110296  
KEYWORDS  
SOURCE  
Neurospora crassa  
Neurospora crassa  
Fukariyora; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Sordariomycetiales; Sordariaceae; Neurospora.

REFERENCE  
1  
AUTHORS  
Mekeyev, E. and Bamford, D.  
TITLE  
Soluble rna polymerase protein and methods for the use thereof  
JOURNAL  
Patent: WO 2004035784-A 3 29-APR-2004;  
RNA-Line Oy (FI)

FEATURES  
Location/Qualifiers  
Source  
1..3078  
/organism="Neurospora crassa"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:5141"

ORIGIN  
Query Match 73.2%; Score 3078; DB 6; Length 3078;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3078; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1129 GCTCGAG 1188  
Db 1 GCTCGAG 60  
Qy 1189 CCGCTGAGAAATTTTGGCCGAATTTCCCAATGAGCTACAGAGAGCTCTCGCTGT 1248  
Db 61 CCGCTGAGAAATTTTGGCCGAATTTCCCAATGAGCTACAGAGAGCTCTCGCTGT 120  
Qy 1249 GCATGGAGAGTTACAGACTCTTTATGCACTGCAAGAGAGAGAGAGAGAGAGAGAG 1308  
Db 121 GCATGGAGAGTTACAGACTCTTTATGCACTGCAAGAGAGAGAGAGAGAGAGAGAG 180  
Qy 1309 GGCCTAAGTAAG 1368  
Db 181 GGCCTAAGTAAG 240  
Qy 1369 CTCTACCGGCTTGATGCTTCCGTGTAACCTTTCCAGAAAGAGAGAGAGAGAGAGAG 1428  
Db 241 CTCTACCGGCTTGATGCTTCCGTGTAACCTTTCCAGAAAGAGAGAGAGAGAGAGAG 300  
Qy 1429 TTGCTAAG 1488  
Db 301 TTGCTAAG 360  
Qy 1489 GTTCTAGACTACAAATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1548  
Db 361 GTTCTAGACTACAAATCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420



QY 1549 CCGCTATGTTGAGCAGAGGCTGCTGACTCAACCGTGGATTGCGTCTGATAGGTTTTC 1608  
DB 421 CCGCTATGTTGAGCAGAGGCTGCTGACTCAACCGTGGATTGCGTCTGATAGGTTTTC 480  
QY 1609 GAGATCTTATACCGTCCCTACGAGCAGCCCAAGTGTACCGCGGTGTGACGAAA 1668  
DB 481 GAGATCTTATACCGTCCCTACGAGCAGCCCAAGTGTACCGCGGTGTGACGAAA 540  
QY 1669 CAACCAAGTGGCTGAGAAAGTCAATCAGTGGCTCAAGATGGGCAACATTTCTTGT 1728  
DB 541 CAACCAAGTGGCTGAGAAAGTCAATCAGTGGCTCAAGATGGGCAACATTTCTTGT 600  
QY 1729 GCGCGCAATGAGCGGCTTTCTTGCGCAAAAGTGGGATACAGGAAACCTCTGAGGAG 1788  
DB 601 GCGCGCAATGAGCGGCTTTCTTGCGCAAAAGTGGGATACAGGAAACCTCTGAGGAG 660  
QY 1789 TTCCAGCTCCGCGCGAGAACCCGAAACCATCATCAAGAGAGATCCTTTTGC 1848  
DB 661 TTCCAGCTCCGCGCGAGAACCCGAAACCATCATCAAGAGAGATCCTTTTGC 720  
QY 1849 GAGACCGGATTAAGTTCGACCTGATGTGTTCAAGAGAGATCTGTCTCCGACAG 1908  
DB 721 GAGACCGGATTAAGTTCGACCTGATGTGTTCAAGAGAGATCTGTCTCCGACAG 780  
QY 1909 GAACCTGAGAGCAACGAGCCGAGTTCAAAGTTAGTCAAAATGCTGGAATGCTGCTGCA 1968  
DB 781 GAACCTGAGAGCAACGAGCCGAGTTCAAAGTTAGTCAAAATGCTGGAATGCTGCTGCA 840  
QY 1969 CTGCAACAACAACCTTGGCAGCCGCACTCAAGTTGTTCTCCGATACAGCTGGCTG 2028  
DB 841 CTGCAACAACAACCTTGGCAGCCGCACTCAAGTTGTTCTCCGATACAGCTGGCTG 900  
QY 2029 AGTAAGACATATGCTATTAATGACATTTGAGGCTCACGATCAACACCAACAACCGAT 2088  
DB 901 AGTAAGACATATGCTATTAATGACATTTGAGGCTCACGATCAACACCAACAACCGAT 960  
QY 2089 CTTCTTCACTTGGCAGCTGGCAGATGATGAAATGAGGCTGTAAGGCGGATGCGGA 2148  
DB 961 CTTCTTCACTTGGCAGCTGGCAGATGATGAAATGAGGCTGTAAGGCGGATGCGGA 1020  
QY 2149 AGCGTGGCAAGAGATACGCGATGTTCTCGGTTTGGTGTATGTGCTCTGCTGTGCA 2208  
DB 1021 AGCGTGGCAAGAGATACGCGATGTTCTCGGTTTGGTGTATGTGCTCTGCTGTGCA 1080  
QY 2209 GGGGGTTTGGTTGGGCGCAAGGAAATGGGTTATCGACGTTGACAGACAGGCGATGAG 2268  
DB 1081 GGGGGTTTGGTTGGGCGCAAGGAAATGGGTTATCGACGTTGACAGACAGGCGATGAG 1140  
QY 2269 GATTGATCGAGACATACCGTCCGAGGCAAGTGGGAAATGCACTTGTGATTAACAT 2328  
DB 1141 GATTGATCGAGACATACCGTCCGAGGCAAGTGGGAAATGCACTTGTGATTAACAT 1200  
QY 2329 CAACGTAACCTCGAAGTCGAGGCGTGTCTGAAGTGAAGTGAAGTGTCTCAACTA 2388  
DB 1201 CAACGTAACCTCGAAGTCGAGGCGTGTCTGAAGTGAAGTGAAGTGTCTCAACTA 1260  
QY 2389 CAGCTGTAACTGTCTGAAAGATAGAGCCAGGCAAGGTGAAGATGCGGCAAGCAATC 2448  
DB 1261 CAGCTGTAACTGTCTGAAAGATAGAGCCAGGCAAGGTGAAGATGCGGCAAGCAATC 1320  
QY 2449 GGTGACCGTCTTATCAAGATTTGCAAGACAGTTCAAGAGCAAAAGCATGTTGAT 2508  
DB 1321 GGTGACCGTCTTATCAAGATTTGCAAGACAGTTCAAGAGCAAAAGCATGTTGAT 1380  
QY 2509 CGCCAGTGAATTTCCGCAATGGGTTTACGAGATTAATTCAGTCCGCAACTCGATC 2568  
DB 1381 CGCCAGTGAATTTCCGCAATGGGTTTACGAGATTAATTCAGTCCGCAACTCGATC 1440  
QY 2569 AGCCAGGCGGTGCTTTTCTTGTGCTGGCTACTGACAGTCAAGAGAGACACTGAAC 2628  
DB 1441 AGCCAGGCGGTGCTTTTCTTGTGCTGGCTACTGACAGTCAAGAGAGACACTGAAC 1500

QY 2629 TTTGTGATGAACAGTGGTTCATCCCAAGAACAAAAGTACTTGAAGACATGCTCG 2688  
DB 1501 TTTGTGATGAACAGTGGTTCATCCCAAGAACAAAAGTACTTGAAGACATGCTCG 1560  
QY 2689 GATCTTCAAAAGCGGAAATGTGACAGTGTGAAGTCCAAAGTGAACATCCGTGCTGA 2748  
DB 1561 GATCTTCAAAAGCGGAAATGTGACAGTGTGAAGTCCAAAGTGAACATCCGTGCTGA 1620  
QY 2749 TCACATATCATTTATCATGATTTGCGGGGTGCTGAGGAAATGAGGTTTAT 2808  
DB 1621 TCACATATCATTTATCATGATTTGCGGGGTGCTGAGGAAATGAGGTTTAT 1660  
QY 2809 GTCCGATTTCTCTCAAAAGTTCAGAGACGAGAGAGATCTTTTACATCTCTAGCTGT 2868  
DB 1681 GTCCGATTTCTCTCAAAAGTTCAGAGACGAGAGAGATCTTTTACATCTCTAGCTGT 1740  
QY 2869 GATGTCTCTGTGGCGATCCAGCCCATTTCCCTAGATATCAACGGGTTGAGCA 2928  
DB 1741 GATGTCTCTGTGGCGATCCAGCCCATTTCCCTAGATATCAACGGGTTGAGCA 1800  
QY 2929 GTCTTCAAGCCAGGCTCCACAGTCTCAAGATGTATCATCTTCTTACTTAAAGAT 2988  
DB 1801 GTCTTCAAGCCAGGCTCCACAGTCTCAAGATGTATCATCTTCTTACTTAAAGAT 1860  
QY 2989 GTACCGCTTGCTAAGAGCTATCTGTGAGACTACAGACGCGATATGCGTGGCTGTC 3048  
DB 1861 GTACCGCTTGCTAAGAGCTATCTGTGAGACTACAGACGCGATATGCGTGGCTGTC 1920  
QY 3049 TGGGATCCGAGATGCTGATGTTTCTGCAATGCGGAAATGCTCTGAGCCGACTG 3108  
DB 1921 TGGGATCCGAGATGCTGATGTTTCTGCAATGCGGAAATGCTCTGAGCCGACTG 1980  
QY 3109 TCTAGTATCTTAAAGAGGACAAACGATTTCAAAACATTAATGCTCTCACAGGACG 3168  
DB 1981 TCTAGTATCTTAAAGAGGACAAACGATTTCAAAACATTAATGCTCTCACAGGACG 2040  
QY 3169 GGCCTCAGCGGCAAGAGACAGACTACATACATATGATCCAGAGACTTCCATTTGCC 3228  
DB 2041 GGCCTCAGCGGCAAGAGACAGACTACATACATATGATCCAGAGACTTCCATTTGCC 2100  
QY 3229 CTGAGCCCAACTTCTTGGGATGTGACTTAACTAACAAAGAGCTCTGTTACATCA 3288  
DB 2101 CTGAGCCCAACTTCTTGGGATGTGACTTAACTAACAAAGAGCTCTGTTACATCA 2160  
QY 3289 AATAGTGTCTTAAACAGCCGCGCATCATTTAGTTCACTGTGGGAAACCTGTGAT 3348  
DB 2161 AATAGTGTCTTAAACAGCCGCGCATCATTTAGTTCACTGTGGGAAACCTGTGAT 2220  
QY 3349 CAGAGCAAGCAAGGATATGTTCTTAAAGAGCAAGTGGCTCAATTTGCGTGAAGACTG 3408  
DB 2221 CAGAGCAAGCAAGGATATGTTCTTAAAGAGCAAGTGGCTCAATTTGCGTGAAGACTG 2280  
QY 3409 CTTGGCGGTGATTTGCTCTTCTGACCCAAATGTACAAAGACGACAGTTGGCTGGGCGC 3468  
DB 2281 CTTGGCGGTGATTTGCTCTTCTGACCCAAATGTACAAAGACGACAGTTGGCTGGGCGC 2340  
QY 3469 GGAAGGCTTACCCCATTTATGATTAAGTGAAGTGTGCTCAATGCGGCAAGCTGCAATGAC 3528  
DB 2341 GGAAGGCTTACCCCATTTATGATTAAGTGAAGTGTGCTCAATGCGGCAAGCTGCAATGAC 2400  
QY 3529 AAGGAATGGAAGGCTTCAACATATGATGAAGAGCGGCAAGATACAGAAAGCGGCGCT 3588  
DB 2401 AAGGAATGGAAGGCTTCAACATATGATGAAGAGCGGCAAGATACAGAAAGCGGCGCT 2460  
QY 3589 CACTTTTGGGATCCGATCTTCTCTTCTTCACTACAGTTTCTTCAAGAGATTAAGGACAG 3648  
DB 2461 CACTTTTGGGATCCGATCTTCTCTTCTTCACTACAGTTTCTTCAAGAGATTAAGGACAG 2520  
QY 3649 TCGGATGCTCCGACGCTATTAAGACACTTGAAGACCGTATCCGCGAAGTGAAGAAA 3708  
DB 2521 TCGGATGCTCCGACGCTATTAAGACACTTGAAGACCGTATCCGCGAAGTGAAGAAA 2560  
QY 3709 GAATATGAGAGTTGTCAAAAACAAAGAGATGAGAGACAGCAAGACCCCTACCTGTG 3768

Db 2581 GAATATGGCAGGTTGTCACAAAACAAAGAGATGAGACAGCAAGACCCCTACCCCTGTC 2640  
QY 3769 CGCGTCAACCAAGGTTTATGAAAAATGTCGCCCATCAAGCCTGAGGCATGACAAATCC 3828  
Db 2641 CGCGTCAACCAAGGTTTATGAAAAATGTCGCCCATCAAGCCTGAGGCATGAGCAAAATCC 2700  
QY 3829 GAGCAAAATATGATTTCTTAAGATGATCAGTTGCTGAGAGCTGTCCTTCGCGGACCGT 3888  
Db 2701 GAGCAAAATATGATTTCTTAAGATGATCAGTTGCTGAGAGCTGTCCTTCGCGGACCGT 2760  
QY 3889 GAGTAAATCATGGGATGCTGAGGGCTAGCAGCGCTTTCAGAGCTGATACCAACAG 3948  
Db 2761 GAGTAAATCATGGGATGCTGAGGGCTAGCAGCGCTTTCAGAGCTGATACCAACAG 2820  
QY 3949 AGCCCCAAGTTGCTGTCGAGATGCGCGGAGACAGACTCGCTGATCATTAAGCGCAGATG 4008  
Db 2821 AGCCCCAAGTTGCTGTCGAGATGCGCGGAGACAGACTCGCTGATCATTAAGCGCAGATG 2880  
QY 4009 ACGAGCAGACCCGGTGAAGCGCCCGCGCTTGATGACCCGCTTCATGATGCGGCTTG 4068  
Db 2881 ACGAGCAGACCCGGTGAAGCGCCCGCGCTTGATGACCCGCTTCATGATGCGGCTTG 2940  
QY 4069 ATGCGGATTAAGAGTTTAAAGACAGTATGTGCGCAAGCGCTGAGGGCGATGATCGGAG 4128  
Db 2941 ATGCGGATTAAGAGTTTAAAGACAGTATGTGCGCAAGCGCTGAGGGCGATGATCGGAG 3000  
QY 4129 TACCGTATCCGAGAGCTCTATGAGTCTGCGCGCATGATATTTTGAATGATGTTTC 4188  
Db 3001 TACCGTATCCGAGAGCTCTATGAGTCTGCGCGCATGATATTTTGAATGATGTTTC 3060  
QY 4189 ACAGGGAATGCGGATTTAT 4206  
Db 3061 ACAGGGAATGCGGATTTAT 3078  
RESULT 5  
LOCUS AF429315 125020 bp DNA linear PRI 18-JAN-2002  
DEFINITION Homo sapiens junctophilin 3 (JPH3) gene, partial cds.  
ACCESSION AF429315  
VERSION AF429315.1 GI:17646244  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1. (bases 1 to 125020)  
Holmes, S.E., O'Hearn, E., Rosenblatt, A., Callahan, C., Hwang, H.S.,  
Ingber, S.L., Ashworth, R.G., Fleisher, A., Stevanin, G., Brice, A.,  
Potter, N.T., Ross, C.A. and Margolis, R.L.  
A repeat expansion in the gene encoding junctophilin-3 is  
associated with Huntington disease-like 2  
Nat. Genet. 29 (4), 377-378 (2001)  
TITLE  
JOURNAL MEDLINE  
PUBMED 21583737  
REFERENCE 2. (bases 1 to 125020)  
Holmes, S.E., Ingber, S.L., Ashworth, R.G., Ross, C.A. and Margolis, R.L.  
Direct Submission  
TITLE Submitted (05-OCT-2001) Psychiatry, Johns Hopkins Medical  
JOURNAL Institutions, 600 N. Wolfe St., Baltimore, MD 21287, USA  
FEATURES  
Source  
1. 125020  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/map="16q24.3; Between D16S520 and WI-12410"  
/note="Isolated from a patient with Huntington's  
Disease-Like 2 (HDL2)"  
complement(35581..35746)  
/rpt\_type=tandem  
/rpt\_unit="ctg"

gene complement(<36507..>36887)  
/gene="JPH3"  
/note="synonym: JP3"  
mRNA complement(<36507..>36887)  
/gene="JPH3"  
/product="Junctophilin 3"  
complement(<36507..36887)  
/gene="JPH3"  
/note="component of the junctional complex between plasma  
membrane and endoplasmic reticulum"  
/codon\_start=1  
/product="Junctophilin 3"  
/protein\_id="AA140941.1"  
/db\_xref="GI:17646245"  
/translation="MSSGRFNPDDGSGYCGGMDKAGHGVCTGPKGGEYTGMS  
HGFVGVYTWPSGNTYQGTMAQGGKNGHIGLESKKMYKJEWTHGFKRGVRECAQ  
NGAKYEGTMSNGLDGVTETYSQD"  
ORIGIN  
Query Match 1.3%; Score 55.2; DB 9; Length 125020;  
Best local similarity 10.6%; Pred. No. 0.01;  
Matches 102; Conservative 397; Mismatches 460; Indels 2; Gaps 1;  
QY 1389 CCCTGTAACCCCTTTCCAGAAAAAGCCGCCCAACGACGTGTTGTAACCGCAATGACGG 1448  
Db 16887 MCYSSKMSMBMSVSVYVMSHMASBSQMBKCMTMSCCMSMYKSSMWGSSWGMC 16946  
QY 1449 CAACCTTGAAGAGAAAGTATGATCCGCTGCTCTGCTGCTTCTAGACTAATCCGA 1508  
Db 16947 WGRRRSKMKTWYRGRMSRSMYTGSGKMRBSMMCTSSCYVSMCMCSCCRSCC 17006  
QY 1509 CAATCGCTTACTCGCCCTTTACTTGTGAAGCTGAAGCCGCTCATGTTGAGCAGAG 1568  
Db 17007 CMCRSYCCMRKYCACACCYMSYWTWASASYSRYSRYSKMSKMSRBSRSC 17066  
QY 1569 CTGTGACTACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1628  
Db 17067 CGSGMGKGGKSGYGRKTKRSGKMGAKMYWYRBSRKMYSKMYCMYCMWGR 17126  
QY 1629 TACGACACCAACCCCAAGTGTACCGCGGTGTCAGCAAAACAGAGTGGCTGCAAGA 1688  
Db 17127 RGCYCGMTSRSMSCSYCAKAKSMCYSYVMSMKGYVYASVYSGMSKXCMGSGTY 17186  
QY 1689 AGTCATCTCACTGCTCAACGATGGGCAACATTTCTTTAGCGCCCAATGCGCGCTT 1748  
Db 17187 STSCGCGCTTTTTCCTCCCAANNTGGGAAGCTTTTNCNKYSYRRKNGCMKYNVYN 17246  
QY 1749 CTTCGCCAAAGATCCGGATACAGAAACCTCTCAGAGAGTCCAGTCCGCGCGAGCA 1808  
Db 17247 NSWRSSCRAGMSTCTKYSMTMSMASYCMWMSYTCMSRMSASRMSWSYMMKMSWR 17306  
QY 1809 CCGAAACCATCATGAAGAGAG--AGTCCATCTCTTTCGCGAGACCGCATTAAGTTTC 1866  
Db 17307 MSYVMSKCMKMYCMMSMRSGMSYVYVMSKSSRGCYTRCYWCMSSSCYSKSY 17366  
QY 1867 CGACTGATGTCTTAAGACGAGATGTCGTTCCGCGAGAGAACCTTAGACAAACGG 1926  
Db 17367 MMRSRKMCMKMSRCSWMSWMSASRSYCYKYSRMSMSKCYRCAGCMWMMKGGYMR 19426  
QY 1927 ACGAGTTCAAGTTAGTAAAGCTGAGCTGCTCTCTGCAACTGCACAAACACTGTG 1986  
Db 17427 YMRCSMKRMMWKSAMRYRRRMRKGRGAMMMCMKYSRMSRMMWKSYSACRSCAMM 17486  
QY 1987 CAGCCGACCTCAAGTTGTTCCCTCATTCAGACTCGTCTGATGAAGATATGCGATT 2046  
Db 17487 SGTVMCASVMCMSSYVCTWCTWSYVSYRCTCWCWGMSSYTSKMSMSSMSYKGRRK 17546  
QY 2047 ATGACATTTGAGGCTTACAGATCAGACACACAGAACCGATTTCTTACCTTAGGC 2106  
Db 17547 SYVMCTSRGAMSCMRCCYMRGASMBRAGSMRBRRAKGRSMGRKMYMTGGMRSXY 17606  
QY 2107 ACTGGGAAGTATGATGACCGTGTAGCCGCGATGTCGGAAGCGTGGCGAAGAGATA 2166



Db 17607 YCTGRHMMTTCMCWRBRSSMYTYSASWAMKRSKSSWGRMMGASASRRCKGASRSSW 17666  
Qy 2167 CGCATGTTCTCGTTGGTGGATGATGTCCTCTCTCTGCAAGGCGGCTTGGTCCGCC 2226  
Db 17667 CSRMKMRGSCWSSKMKGSSRSASCKGSGRMRRSSSKYRRGKGRKSMTKS 17726  
Qy 2227 AAGGGAATGCGGTTATGACGCTTGACACACACGCGCATGATGATGATGACATAC 2286  
Db 17727 KGSKGRKMKRGRSSMTSSCTYYASSCMMMSKSCMCCMMKRCACCTSSMCTSM 17786  
Qy 2287 CCGTCCAGCGCAAGTGGGAATGCGACTTCTGTATTAACATCAAGTACCTCGAATC 2346  
Db 17787 YRCGWMKSYSTGCTTCKCTGYKKSRTWYMYMTSKMYMRAACMYCYMSTRBM 17846  
Qy 2347 C 2347  
Db 17847 C 17847

RESULT 6  
AF468822 6688 bp DNA linear PLN 14-FEB-2002  
LOCUS Diaporthe perijuncta putative RNA-dependent RNA polymerase RDP-1  
DEFINITION (rdp-1) gene, complete cds; and unknown genes.  
ACCESSION AF468822  
VERSION AF468822.1 GI:18656896  
KEYWORDS  
SOURCE Diaporthe perijuncta  
ORGANISM Diaporthe perijuncta  
Bakaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Sordariomycetidae; Diaporthales; Valsaceae; Diaporthe.  
1 (bases 1 to 6688)  
Preisig, O. and Wingfield, M.J.  
Putative fungal RNA-dependent RNA polymerase  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 6688)  
Preisig, O. and Wingfield, M.J.  
Direct Submission  
Submitted (15-JUN-2002) Forestry and Agricultural Biotechnology  
Institute, University of Pretoria, Lumnon Street 74, Pretoria 0002,  
South Africa  
FEATURES  
source Location/Qualifiers  
1..6688  
/organism="Diaporthe perijuncta"  
/mol\_type="genomic DNA"  
/isolate="CMW3407"  
/db\_xref="taxon:186170"  
/country="South Africa"  
join(<1..275,348..>381)  
/product="unknown"  
join(<1..275,348..>381)  
/note="similar to D-lactate dehydrogenase (D-lactate  
fericytochrome C oxidoreductase)"  
/codon\_start=1  
/product="unknown"  
/protein\_id="AAL78033.1"  
/db\_xref="GI:18656897"  
/translation="MDDLGLFASILGHIGDNFHESLILYNPVGNERKEVECVKMYR  
RAIMDGTCTGSEHSIGWKKDLSLLEVGPELTAVMRAIKGSLDPRMIMNPKIPDRS  
"  
" <1361..>5870  
/gene="rdp-1"  
/gene="rdp-1"  
join(<1361..4598,4654..>5870)  
/product="putative RNA-dependent RNA polymerase RDP-1"  
join(<1361..4598,4654..>5870)  
/gene="rdp-1"  
/codon\_start=1  
/product="putative RNA-dependent RNA polymerase RDP-1"  
/protein\_id="AAL78034.1"  
/db\_xref="GI:18656898"  
/translation="MSHHSPTPTPDGSPSAQPDASAQIHHTTQTSIPDATAKDNPR  
ISTLRPAIADTSDASNRNTPSVRLPSPPLSMSTHSNGNVILGQKVSLSGPIQGRS

IOINARTFEGPHGHPNPRONRPNRROORSPOSILQORSTPAITGAVLSAQSILTI  
DLFGLPTDIKIREITWESMTSNLNVNRVEIYENFDKRGKAKVIEIPPIRMPYMS  
PNLITIDBADSSLOVETREPPFSTETPCNRPIPAORCOKKILPTLVFPMMAE  
SAMLNPFORSYEDVVLITTRSKROFLOQTIVNIGRGCAFKSQIDPSHTKIKMR  
ENKAMLAHLPLPPLYLRLPVSHDEDTLMOMBLMAQOTISSMHLRLPALU  
EKOTWSIEFRWLTYYLDMESLSMESRYNTLADLVNKLILQVHDFMLNLSNM  
KLIDNAGEVTEILGTSQYHLPVRYOLEVCSISGLSEYVNAQLEKLSIFED  
RADMLEAVTEANKOFFEPEMPEMDPRILHYWNNAIIPYATLVRAVILPTIYFKT  
PCVELNRLIRKYSDDLNDRLRYOQTFDETFCKIPSSODSNKDNITRYHRYMONG  
ITIDRHYKTLAPNSOPRENGAFPCFETDHTVTCASIRMMKGDPRHRSVCKFAABMO  
CFTTTRQVNGISIPKIRQIGDIQRSGDNINFTDVGKISVFPAHILASERLPEP  
SCFQMMGCKGKVYVVPDIPIANVHIRSOEKFVAVNGLEIKSTFPAHLNKOY  
IPVLILAGVDAVFVMDLDELKEYEALDSKASELLRSQDEMOQTLTYMNEJDT  
FMDSPEPMTLRLMKCWILKELKOKKPAISVSGNSMIKGVNDETVGLRHSQDTGK  
GPNITSLPOITFLOVPIEGSDKSTNVEYITGICVGNRPALHPDQVRYEAVDPE  
LRHLKVVVFPKTKGDRDVPSCGSGMDGDVYVYVYDERLIPSEMPHPLDHDADST  
PGIDSPADVTNGDVTRFAOYMNDSLGRILATIRFADQDLAGVAPKTEILAKHS  
KAVDYIKSGPVMKRLQPRMNMHMERDKSSYHSGALQOIVRIKIEFPAAE  
MSFOTRILSRYOLEDLTLAASEIKAAVYIAMRLTGOHAPVTEFIEWTFILYKPR  
VGSDFVLOENVGRMAALKARPAVCAVEATGDEOTPTCASASSKVLELDFVAM  
YTYVHNEVRAAVRESRPRPDGDLLEIOMPLISFPMLYHRELRYALGGDVRPL  
RKPVCAADPAANIESPESGDALSEBIDQEGTOLSLQKRSKDKQYVIDRGSVERG  
KVGEDYVRTSSGVVRGDLTLFTDGEVQAGSTIAFATPAENNSNPLVEDDG  
TGSYSYPASPRSEHEABGEDEIEFBEVDGANDBEEDALEILARKIGM"  
<6068..>6526  
/product="unknown"  
6068..6526  
/codon\_start=1  
/product="unknown"  
/protein\_id="AAL78035.1"  
/db\_xref="GI:18656899"  
/translation="MTLONHQANRPPLSLITYSSRAVDYLISLVNMGDCFPITWGP  
MMNQPADDLGTSPTGTSRDPDPWHQQRDSQHAFLQLAGTSLDYLVDWIGLSAEIP  
ARLYSSLSIHLRPSRSTMDSRVSSSLRTADDSQKVAHQCLFPLP"

ORIGIN  
Query Match 1.3%; Score 53.8; DB 8; Length 6688;  
Best Local Similarity 54.9%; Pred. No. 0.024;  
Matches 106; Conservative 0; Mismatches 87; Indels 0; Gaps 0;  
Qy 2874 CTTGTGGCGCATCCCGACCCATTTCTCTAGTATATCAACGGTTGACGACTT 2913  
Db 4207 CGTAGTAGCCGTAACCCACGCGCTTCTCTGAGATGTCGGGTTGTTAGCGATTGA 4266  
Qy 2934 CAAGCGAGACTCCACAGTCTCAAGATGTAATCATCTTCTACTAAGAAGATGATACC 2993  
Db 4267 CGTGCCAGAGTTACGACATCTCAAGACGATGCTGCTTCCAAAGCTGAGACAGGA 4326  
Qy 2994 GCTTGCTAAGAGCTATCTGTGAGACTAGACGCGGATATGAGCTGTGCTGGGA 3053  
Db 4327 CGTTCGACATGTGCTCCGATGATGATGACGAGATGACTACTTGTGTTACTGGGA 4386  
Qy 3054 TCCGAGATCGTC 3066  
Db 4387 CGAGCGACTCATC 4399

RESULT 7  
AX655393 2000 bp DNA linear PAT 22-MAR-2003  
LOCUS Oryza sativa  
DEFINITION Sequence 5263 from Patent WO03000898.  
ACCESSION AX655393  
VERSION AX655393.1 GI:29158207  
KEYWORDS  
SOURCE Oryza sativa  
ORGANISM Oryza sativa  
Bakaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
Ehharitoidae; Oryzaceae; Oryza.  
REFERENCE 1  
AUTHORS Chang, H. S., Chen, W., Cooper, B., Glazebrook, J., Goff, S. A., Hou, Y. M.,  
Katagiri, F., Qian, S., Tao, Y., Whitham, S., Xie, Z., Zhu, T., and Zou, G.  
TITLE Plant genes involved in defense against pathogens



1991 CGCAGCTCAAGTTGTTCTCCGATACAGCTCGCTGCTAGTAAGACATATGCACTTATGA 2050  
172 YGAMCMSCSMNMTMSGCGCTTGMRKSKYSCCKRYCCCTKYCTGYRYCKMYK 113  
2051 CATTGAGCCTCACCAGATCAGACACCAAGACGATCTTCTTACCTTCAGCACTG 2110  
112 YSYKCYCYCYWYMYRMMKCMSCSMMSCATCTSTSRMMWYTYAAMKMGCG 53  
2111 GCGAAGTCATGAATGACGCTGAG 2135  
52 SSGMYRMSKSKMYSKYSSCKYTS 28

RESULT 9  
LOCUS A10243 1164 bp DNA linear PAT 25-JAN-1994  
DEFINITION S. murinus DNA for glucose isomerase.  
ACCESSION A10243  
VERSION A10243.1 GI:490202  
KEYWORDS glucose isomerase.  
SOURCE Streptomyces murinus  
ORGANISM Streptomyces murinus  
Bacteria; Actinobacteridae; Actinomycetales;  
Streptomycinae; Streptomycetaceae; Streptomyces.  
REFERENCE 1 (bases 1 to 1164)  
Luiten, R.G.M., Quax, W.J., Schuurhuizen, P.W. and Mabel, N.  
Novel glucose isomerase enzymes and their use  
Patent: EP 0351029-A 16 17-JAN-1990;  
GIST-BROCADES N.V.; PLANT GENETIC SYSTEMS, N.V.  
FEATURES  
source location/Qualifiers  
1..1164  
/organism="Streptomyces murinus"  
/mol\_type="unasigned DNA"  
/db\_xref="taxon:33900"  
1..>1164  
/codon\_start=1  
/transl\_table=1  
/product="glucose isomerase"  
/protein\_id="CA00885.1"  
/db\_xref="GI:490203"  
/db\_xref="GOA:P37031"  
/db\_xref="UniProt/Swiss-Prot:P37031"  
/translation="MSTQPTPEPRFTREGLMTVNGQRPDGDATRPALDPVETVORLA  
ELGAYCTFHDDLIIPRESSPTRESHIKRRQALDAGTVVPAATTLFTRPFXKG  
GFYANDRVRKYLKTTIGNIDLAELGAKTYVAMGSESGAKDVRALDRMKE  
AFDLIGEYVTAQGYDLRPAIEPKNEPRGDIILPTVGHALFIERLSRPELVNPEV  
GHEOMAGLNFPHGIAQALMAGLKFHIDNGSGIKYDDLRFAGDLSRAAFVLDLLE  
TAGEGPRHFPKPRTEPDFGVWASAGCKRNYLLIKDRRAAFRADEVEALRAAR  
LDQLAQPRAAAGDLALADRAAFDFVDAAAGMAFEHLIDLAMHLGARG"

ORIGIN  
Query Match 1.2%; Score 51.6; DB 6; Length 1164;  
Best Local Similarity 47.8%; Pred. No. 0.089;  
Matches 150; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

256 GTGCCATGGCAGACCAACCACTTCAACCAACCCGCTTCAAGAGACGGCGCTTACC 315  
331 GCTACTCGCGCCA 344  
316 GCCAACGACCGCA 329

RESULT 10  
LOCUS AF443073 3524 bp DNA linear PLN 03-DEC-2001  
DEFINITION Phomopsis sp. CMW 5588 putative RNA-dependent RNA polymerase  
(rdp-1) gene, partial cds.  
ACCESSION AF443073  
VERSION AF443073.1 GI:17227101  
KEYWORDS  
SOURCE Phomopsis sp. CMW 5588  
Phomopsis sp. CMW 5588  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Sordariomycetidae; Diaporthales; Valsaceae; mitosporic Valsaceae;  
Phomopsis.  
REFERENCE 1 (bases 1 to 3524)  
Preisig, O. and Wingfield, M.J.  
Direct Submission  
Submitted (04-NOV-2001) Forestry and Agricultural Biotechnology  
Institute, University of Pretoria, 72 Lunnon Street, Pretoria 0002,  
South Africa  
FEATURES  
source location/Qualifiers  
1..3524  
/organism="Phomopsis sp. CMW 5588"  
/mol\_type="genomic DNA"  
/isolate="CMW 5588"  
/db\_xref="taxon:179291"  
/country="South Africa"  
<1..>3524  
/gene="rdp-1"  
/join(<1..2092,2148..>3424)  
/gene="rdp-1"  
/product="putative RNA-dependent RNA polymerase"  
/join(<1..2092,2148..>3424)  
/gene="rdp-1"  
/note="rdp-1"  
/codon\_start=1  
/product="putative RNA-dependent RNA polymerase"  
/protein\_id="AAJ2801.1"  
/db\_xref="GI:17227102"  
/translation="HKLKHLPSLVKPVQGVSEFRWTTYYFDINSSLSSESVRT  
YLADYVKTQVNRBELVTPGSPNANMLDNSAEMTLELSSTYTLDFQVYQLFV  
CISGLDDEYSVNAQFLDKNSPDADARMLGVAEASIQLEHPMKI PDDPKVLYH  
PNARVPYATVLRRAVITPTTIVKTPCVETLRILRKYSDDLNDRLPLRQFCEITFG  
KITSSQDSKNDMLYIRVHVMONGISIGRHYRPLAFNSQRENGAFPCETDHYT  
CDSTRMWMGDPRHRSVGRFAARMGOCFTTRVNGISITNIRKIDIEBQOTHTN  
FTPOVGKISNFKLASERDLPEBSCFQMRIGCGGVLVWVPDVPSEVHVRPSQ  
KFXRAYVNLIIKTFAFSATFPNKQVLPVLAAGVDSVSVMSLMDLREYDEALDS  
MKGELLRSOVDENOTLTMAEMVDTEMDKEPFLWTLRLMKMCLIKRLKHFALIS  
KSGAMVGVVDIEIGVLRGHSOATREGVNSIESLPOIFLQVPIEGSGRSTNYEVT  
GICVYGRNPSLPHEDVRVYZAVDYPELRHLYKVVYFPEKTDROI PMSGSGDNDGY  
FVWDERLITPEWDHPPLNHDAGSSTLDKPADVTEBVTFFQYMKNDLSLGRATAM  
FAVDQSLSGGVNPKCIELAKLSMAVDYIKSGRPAVMKHLDPKPKPMMEKESN  
YRSYSAIGKIYDRIKIEEPFAAYEMPDAIILSRVQLEADTLAKAKIKATYDANMR  
LMGOSHAEPVNEFIEWSTFELSRPVGSDYDLQSVLEALALKLFABGCEAVTANR  
OTGCFAPASSVNLKLDKRVAMVYTHNDVAAALBERMPKNGSGSBEIQMPI  
SPPLFTRIELARVALGRGVYRPLQKSVYMLNRGESOOTRQDVAVEAEFIIIDPTL  
PDALDREKTOQAVGDVQSKEGTTERGSGADGEIGDCCVRTSTQVVRHGMML  
FTETGEVQGOQSRMPTDISPAGTQGIQSASVNSVLVADSTGSSCPASPSRERE  
VVEEBEDIEVEVDGAYEEVEDALELARKLGI"

ORIGIN  
Query Match 1.2%; Score 51.4; DB 8; Length 3524;  
Best Local Similarity 55.2%; Pred. No. 0.1;  
Matches 100; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

Db	1707	CGTGGTGGCCGCATCCATCCATCCATCCGAGATGTTACGAGTGGTCGAGGCAGTGA	1766
Qy	2934	CAAGCAGAGCTCCACAGCTCTCAAGAGATGTAATCATCTTCTCTACTTAAGGAGATGTACC	29933
Db	1767	CGTGCGCTGAGCTTCGACATCTGAGAAAGATCGTGGTGGTCTTCCCAAGACTGGCGACAGAA	1826
Qy	2994	GCTTGCTAAGAGCATATGTGGTGGAGATACGACGGCGATATGGCTGGTCTGCTGGGA	30533
Db	1827	CATTCGCCAGCATGTGTCTGGAGGTGATGACGAGGATGATCATTTGTCTATTGGGA	1886
Oy	3054	T 3054	
Db	1887	T 1887	

RESULT 11

AF429315/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

repeat\_region

gene

mRNA

CDS

ORIGIN

Query Match

1.2%; Score 50.8; DB 9; Length 125020;

	Best Local Similarity	11.5%	Pred. No. 0.15;	Matches 110;	Conservative 407;	Mismatches 428;	Indels 9;	Gaps 1;
QY	1698	GTGGCTCAGATGGGGCAACATTTCTCTTTGATGCGCCGCAATGCGCGCTTTCTTTCGCCAA	1757					
Db	51541	GAGCCSAGCGTGTGTSRMSACCKYMSKKSARARVAYASTGAAMSMCGTCYMKQCTR	5148					
QY	1758	AGATCCGGGATPACAGAAACCTCTCAGGAGTTCCAGCTCCGCCGACGACCCGAAAC	1817					
Db	51481	MMATACMRRRARMAWMSCKMRKCTWSKSRGRRMRRCWMSARWMYAAGYSMKCSRWG	51422					
QY	1818	CATCATCAAGAGAGAGTCCACTTCTTTGGCCGAGCCGGATTACGTTACGTCGACTGT	1877					
Db	51421	GYSRSRSRSGMSRRTGYSASCTYMSKAGAGKCKGASTAKSGAGRMSYTKGAW	5136					
QY	1878	GTTCCAGACGAGATCTGCTCGTCCGACGAGAAACCTGTAGACACGACCGAGTTCAA	1937					
Db	51361	YSCMSASKKSRLMTCVSCCCTSGCKGYMYMYMKYAGCWCWYMRNCTCTMAARARAR	5130					
QY	1938	AGTTAGTCAATGCTGAGTGGCTCTGCTCACTCGACACACACTTGGACCGCGACCT	1997					
Db	51301	AAMRRARWAGARARWMSMMWMMBHAAAMARWMAAKCMWYYSYRKYCMSWMA	51244					
QY	1998	CAA-----GTTGTTCCCGTATCCAGCTCGCTGTAGTAAGACATATCCATTAT	2048					
Db	51241	SWAAMAMMMCTAAWYKKWTRMRMSKYSSTVAAYATWCYCYSSMTAARMGTGSG	5118					
QY	2049	GACATTTGAGCGCTCACCGATTCAGACACACAAAGACCGATCTTCTTACCTTCAGACAC	2108					
Db	51181	GKRRKMASTSCMCCRRGCCMKKKMKMNCSTYTTGMSGGSRMTWSMCMYVYKSRK	51122					
QY	2109	TGGCGAAGTAGTAATGACGCTGTAGGCGCATGTGCGGAAGCGTGGCCAAAGAGATACG	2168					
Db	51121	TKSMATGSCMRMGAMRSKGMGRKYWSCRYKGRMRGMSYKGCSGYMSMGMSGRTSKY	5106					
QY	2169	CGATGTTCTCGGTTGGGTATGTGCCCTCTGCTGTGCAAGGGCGTTTGTTGGCCCA	2228					
Db	51061	CSYTGVCYSKASAMCKSKSTSRCKTKRYSCSMGYSSTCTCWGGRMKCKMKSACAS	5100					
QY	2229	GGAATGATGGGTTATTCACGTTTACGACACAGGCGATGAGATTTGGATCGACATACCC	2288					
Db	51001	RSCKSKYKMASWRCRBRGMSCHYMSMSYSMASTKKSARGRMKCYTSCMKTYYSM	5094					
QY	2289	GTCCACGCGCAAGTGGGAATGCGACTTCGTTGATTAACATCAACGTAACCTTCGAGTCG	2348					
Db	50941	HBKSHDBSGYSKRSYBYWTGMMKMSKMSWSDHMMRDMMYDKRMKYSSTGRYHNB	5088					
QY	2349	GAGCGTGGCTTCTTAATGTAAGTACGCTGGTCTTCAACCTACAGCTGTACTCTCTGGA	2408					
Db	50881	SAWCKMSRSDSGYSGMYHMSKYSMSKSMYYSGMKSSYVBDMCMSMYTBSKSVR	5082					
QY	2409	AGATAGAGCCGAGGACAAAGGTAAGATGGCCACGAGCAATGCGTGAACGCTTTATCAACA	2468					
Db	50821	MRRWSSGYBYRKCAYMRIRYBHSSTRSYMGKSSCWVRDVSRRRSVHTMSMRMC	5076					
QY	2469	TTTGCAACGACAGTTCCAGCGACAAAGACATGCTTTGAATCGCCACAGTGAATTTGCGCA	2528					
Db	50761	RCRABARWAKSAGKMSCVMTGMSCHMSKMSKMSWDDCCHABADHAGVRCRWKBCA	5070					
QY	2529	ATGGGTTTACAGAGTTATTTCCAGTGGCGCACTCGAGTCAGCCAGCGCCGTGCTCTTT	2588					
Db	50701	DBCYKRDYSVSCWSAHVYVWVBHBSKSSRBGMRHMHMACHKMSCKKSUVVWVHMKKCY	5064					
QY	2589	TCTTGCTGGGCTACCTACAGTCAAGAGAGACACTGAATTTTGTATGAACAG	2642					
Db	50641	WCRRSCWBCSNGVHGRDVRVKKRTAAGGGRARMYKRAKAYDDSSDMWSWG	5058					

ACCESSION AB038265.1 GI:14669805  
VERSION  
KEYWORDS D-xylose isomerase.  
SOURCE Streptomyces olivaceoviridis  
ORGANISM Streptomyces olivaceoviridis  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Streptomycinae; Streptomycetaceae; Streptomyces.

REFERENCE  
1 Kaneko, T.  
TITLE Streptomyces olivaceoviridis D-xylose isomerase (xy1A) gene  
JOURNAL Published Only in Database (2001)  
AUTHORS  
2 (bases 1 to 1610)  
TITLE Kaneko, T.  
JOURNAL Direct Submission  
Submitted (10-FEB-2000) Takanito Kaneko, Akita Research Institute  
of Food and Brewing, Bioengineering Div.; 4-26 Sanuki, Araya-machi,  
Akita, Akita 010-1623, Japan  
(E-mail:kaneko@ake.arif.pref.akita.jp, Tel:+81-18-888-2001)  
location/Qualifiers  
1. 1610  
/organism="Streptomyces olivaceoviridis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:1921"  
268..1434  
/gene="xy1A"  
268..1434  
/gene="xy1A"  
/codon\_start=1  
/transl\_table=1  
/product="D-xylose isomerase"  
/protein\_id="BAB62014.1"  
/db\_xref="GI:14669805"  
/translation="MSYQPTPEDRFTFGILMTVGMSQGRDPFGDATTRALDPEVYQRLA  
ELGAGVTFHDDLIIPGSSDTERESHIKFRQALDAGTVPATNLTLPVFKG  
AFANDVDVRYALRKTIRINDLAELGATYVAMGREGAEGSAADVSLDRME  
AFDLGIVYTSQGVLFPALEPKNEPRGDIILPTVGHALAFIRLERPELVGNPEV  
GHOMAGLNPFGIAQALWAGKLFHIDNGSGIKYQODLRFAGDLSALMVDLLE  
SAGEGPRHDFRPTREDIDGVMAAGCMRYLILKEAAAARADPVOALRASR  
LDQLAQTAAADGLELDLADRAAFEDPVEAAAAGMAFERLDQALMHLGARG"

ORIGIN  
Query Match 1.2%; Score 50; DB 1; Length 1610;  
Best Local Similarity 47.5%; Pred. No. 0.23;  
Matches 149; Conservative 0; Mismatches 165; Indels 0; Gaps 0;  
QY 31 AGCCCGTCGAGAAATCATTAACCGGCTCAATTAACGACTCAACCTGGGCTTCAAGTGT 90  
DB 283 ACCCCGAGGACAGGTTCACTTGGACTTGGACCGTGGCAAGGGAAGGACCCG 342  
QY 91 GTGCGACAGCAACTTCAACCCCGCCCGGGAAGAGCTGGCCGAGATGAGAGAT 150  
DB 343 TTCGGCGACCGCACCGCGCGCCCTCGACCGGTCGAGCGGTGACCGCTGGCGAA 402  
QY 151 TTCGGTGGCATGACAAGATCTACAGAGCCCTGAACCTTCTACTGGCGGAAGATGAC 210  
DB 403 CTCGGTGGCCACGAGTGAACCTTCCAGACGACGACCTGATCCCTTCGGTTCGGAC 462  
QY 211 TCCTGGAACAGGAGCAACTTCTTCAATGAGCCAAAGCTGCGAGCTCGAATGG 270  
DB 463 ACCGAGCGCGATGACATCAAGCGGTTCCGCGAGGCGCTGGACGCGACGCGATGACC 522  
QY 271 GTGGCCAAAGCCAGCGCGACCTGACGCTTCCGTGTCCAGAGAACCTCCCGCGCC 330  
DB 523 GTTCGAGTGGCCACGAACTTTCACGACCCCGTTCAGAGAGCGCGCTTACG 582  
QY 331 GCTACTGCGGCGCA 344  
DB 583 GCCAACGACCGGA 596

RESULT 13  
AF170068 1952 bp DNA linear BCT 12-AUG-1999  
LOCUS

DEFINITION Streptomyces chibaensis D-xylose isomerase (xy1A) gene, complete  
cds.  
ACCESSION AF170068  
VERSION AF170068.1 GI:5731284  
KEYWORDS  
SOURCE Streptomyces chibaensis  
ORGANISM Streptomyces chibaensis  
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;  
Streptomycinae; Streptomycetaceae; Streptomyces.

REFERENCE  
1 (bases 1 to 1952)  
AUTHORS Joo, G.-J., Shin, J.H., Heo, G.-Y., Park, H.D. and Rhee, I.K.  
TITLE Streptomyces chibaensis J-59 xy1A  
JOURNAL Unpublished  
2 (bases 1 to 1952)  
AUTHORS Joo, G.-J., Shin, J.H., Heo, G.-Y., Park, H.D. and Rhee, I.K.  
TITLE Direct Submission  
Submitted (19-JUL-1999) Department of Agricultural Chemistry,  
College of Agriculture, Kyungpook National University, Buk-gu,  
Sankyuk-dong, 1372, Taegu 702-701, Korea  
location/Qualifiers  
1. 1952  
/organism="Streptomyces chibaensis"  
/mol\_type="genomic DNA"  
/strain="J-59"  
/db\_xref="taxon:67286"  
552..1718  
/gene="xy1A"  
552..1718  
/gene="xy1A"  
/codon\_start=1  
/transl\_table=1  
/product="D-xylose isomerase"  
/protein\_id="AAD48850.1"  
/db\_xref="GI:5731285"  
/translation="MSYQPTPEDRFTFGILMTVGMSQGRDPFGDATTRALDPEVYQRLA  
ELGAGVTFHDDLIIPGSSDTERESHIKFRQALDAGTVPATNLTLPVFKG  
AFANDVDVRYALRKTIRINDLAELGATYVAMGREGAEGSAADVSLDRME  
AFDLGIVYTSQGVLFPALEPKNEPRGDIILPTVGHALAFIRLERPELVGNPEV  
GHOMAGLNPFGIAQALWAGKLFHIDNGSGIKYQODLRFAGDLSALMVDLLE  
SAGEGPRHDFRPTREDIDGVMAAGCMRYLILKEAAAARADPVOALRASR  
LDQLAQTAAADGLELDLADRAAFEDPVEAAAAGMAFERLDQALMHLGARG"

ORIGIN  
Query Match 1.2%; Score 50; DB 1; Length 1952;  
Best Local Similarity 47.5%; Pred. No. 0.23;  
Matches 149; Conservative 0; Mismatches 165; Indels 0; Gaps 0;  
QY 31 AGCCCGTCGAGAAATCATTAACCGGCTCAATTAACGACTCAACCTGGGCTTCAAGTGT 90  
DB 567 ACCCCGAGGACAGGTTCACTTGGACTTGGACCGTGGCAAGGGAAGGACCCG 626  
QY 91 GTGCGACAGCAACTTCAACCCCGCCCGGGAAGAGCTGGCCGAGATGAGAGAT 150  
DB 627 TTCGGCGACCGCACCGCGCGCCCTCGACCGGTCGAGCGGTGACCGCTGGCGAA 686  
QY 151 TTCGGTGGCATGACAAGATCTACAGAGCCCTGAACCTTCTACTGGCGGAAGATGAC 210  
DB 687 CTCGGTGGCCACGAGTGAACCTTCCAGACGACGACCTGATCCCTTCGGTTCGGAC 746  
QY 211 TCCTGGAACAGGAGCAACTTCTTCAATGAGCCAAAGCTGCGAGCTCGAATGG 270  
DB 747 ACCGAGCGCGATGACATCAAGCGGTTCCGCGAGGCGCTGGACGCGACGATGACC 806  
QY 271 GTGGCCAAAGCCAGCGCGACCTGACGCTTCCGTGTCCAGAGAACCTCCCGCGCC 330  
DB 807 GTCCGATGGCCACGAACTTTCACGACCCCGTTCAGAGAGCGCGCTTACG 866  
QY 331 GCTACTGCGGCGCA 344  
DB 867 GCCAACGACCGGA 880

RESULT 14

ARI45904  
LOCUS ARI45904 3731 bp DNA linear PAT 08-AUG-2001  
DEFINITION Sequence 1 from patent US 6218142.  
ACCESSION ARI45904  
VERSION ARI45904.1 GI:15109093  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
AUTHORS 1 (bases 1 to 3731)  
TITLE Unclassified.  
Massegger M., Riedel, L., Schiebel, W. and Sanger, H. L.  
Nucleic acid molecules encoding polypeptides having the enzymatic  
activity of an RNA-directed RNA polymerase (RDRP)  
Patent: US 6218142-A 17-APR-2001;  
JOURNAL Location/Qualifiers  
FEATURES  
source 1..3731  
/organism="Unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Query Match 1.2%; Score 49; DB 6; Length 3731;  
Best Local Similarity 51.6%; Pred. No. 0.43;  
Matches 112; Conservative 0; Mismatches 105; Indels 0; Gaps 0;  
QY 2850 TACACTCTATCGAGCTGATGCTCTCGTGGCGCGATCCCGACCCATTCCCTAGTGA 2909  
DB 2425 TAAATTCATTCTGAAGGAAATGTGTTGTCAGAAAATCCATGCTTGATCCTGTGTA 2484  
QY 2910 TATCCAAGGGTTGAGCAGCTTCAAGCAGAGCTCCACAGTCTCAAGAGTGAATCAT 2969  
DB 2485 TATTCGTTTAAAGCTGTAATGTTTCAGAGCCCTGCACACATGTAATTTGTGTGT 2544  
QY 2970 CTTCCTACTAAAGAGATGATCGCTTGTCTAAGACCTATCTGTGGAGACTAAGACGG 3029  
DB 2545 ATTCCTCAGAAAAGAAAAGACCTCATCCGAATGATGTTTGGAGTGTATTTGATGG 2604  
QY 3030 CGATATGGCTGGGTCTGCTGGGATCCGAGATCGTC 3066  
DB 2605 GGATATCTACTTTGTTGCTGGATCAAGACATGATC 2641  
RESULT 15  
LERDRP  
LOCUS LERDRP 3731 bp mRNA linear PLN 18-DEC-1998  
DEFINITION L.esculentum mRNA for RNA-directed RNA polymerase.  
ACCESSION Y10403  
VERSION Y10403.1 GI:4038591  
KEYWORDS RDRP gene; RNA-directed RNA polymerase.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
REFERENCE  
AUTHORS 1  
Schiebel, W., Pelissier, T., Riedel, L., Thalmair, S., Schiebel, R.,  
Kemper, D., Lottspeich, F., Sanger, H. L. and Massegger, M.  
Isolation of an RNA-directed RNA polymerase-specific cDNA clone  
from tomato  
JOURNAL Plant Cell 10 (12), 2087-2101 (1998)  
MEDLINE 99055198  
PUBMED 9836747  
REFERENCE 2 (bases 1 to 3731)  
AUTHORS Massegger, M.  
TITLE Direct Submission  
JOURNAL Submitted (08-JAN-1997) M. Massegger, Max-Planck-inst. fuer  
Biochemie, Viroidforschung, Am Klopferspitz 18a,  
Planegg-Martinsried, 82152, FRG  
FEATURES  
source 1..3731  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/strain="Rutgers"  
/db\_xref="taxon:4081"

gene 1..3731  
/gene="RDRP"  
CDS 194..3538  
/gene="RDRP"  
/BC\_number="2.7.7.48"  
/codon\_start=1  
/product="RNA-directed RNA polymerase"  
/protein\_id="CAA71421.1"  
/db\_xref="GI:4038592"  
/db\_xref="GOA:Q9ZR58"  
/db\_xref="UniProt/TREMBL:Q9ZR58"  
/translation="MKTIQVRGFYLLAEVYKSLKTYGTYCALERKSGGS  
RAKAKQFADNISADKIITLANNRLVFGSSYLKAWMKTDIVQLRAYVQMDITLNF  
GCOISDDKPAVLGSTEVSIOFGIKKPFEPFLSSGADYKLQSYENIQVILHRYG  
ONAOFLIOLFGAPRIYKRLNENSCYFPKPEPDQDQVRRVTDPPSPWIGSSLCQFR  
RGVRLPNFESFPHVAERENNTILOTGTFEPFPOKALPNVQPPGISIPYILKPK  
SLVHGCTPGPALVYFPRIVDPRRRNACIEHLAEKYYIKCCYDPRMTTBOYD  
GYLKGQPKPSITLDDGLVYRVLVTPCKVYFCGPEVANSRVLKRYSEDIENFL  
RVSEVDEWEKLYSTDLLPKASTGSGVRTNIYERILSTLRKGFVIGDKKPEFLAFSS  
QLRDNVMMFASRPGLJLANDIRAMQDFSOIKNVAKYARLQSPGSSRETLVLHE  
IEVTPVKYHGSYVPSDGIKISGDFARVASKCGIQTSPSAFOIRYGVKGVAVD  
PDSMKLSLRKMSKYSNDIKLDVGMKGYOCYANROLITLSTLGVKDEVLQKO  
KEAVDOLDLIHDLSLKAQBALMSPGENTILKMLNGCYRPAEPPLSMIOTFRA  
SKLDDRTSRIFIPNGRTMGCIDBSRFLBEGVYVQTGACGEFSDDLHPNRS  
STNSNFI LGNVVAVAKNPLHREDIRVLAVVRAVLAHHVDVCFEYQKRRPHPNCS  
GSDLDGDIYFCWMDQMIIPRQVQPMVEYPPASIQLDHDTVTEEVEYPTNYVNDL  
GIIANHVVFAADREPMDMSDPCKLAEFLAVVFPKPAVAELPSQURPKXPYDFM  
DKPDKTSYISERVIGLFRKVDKAPQASSIATPTFDVARRSYDADAEVDFEDYIDE  
APFKTEYNKSGNLMDYVIGKTEAEILSGGIMKSKTDRDRDAEAIIVAAVALKE  
APAFKRRNDIDIMLPKASAMHYVTHPTPYMGCIYNOGLRAHFIISPMCVYDOLIOIK  
KDKARRPVLNLSLRQLSHRLVLEK"  
ORIGIN  
Query Match 1.2%; Score 49; DB 8; Length 3731;  
Best Local Similarity 51.6%; Pred. No. 0.43;  
Matches 112; Conservative 0; Mismatches 105; Indels 0; Gaps 0;  
QY 2850 TACACTCTATCGAGCTGATGCTCTGTCGGCGAGTCCCGACCCATTCCCTAGTGA 2909  
DB 2425 TAAATTCATTCTGAAGGAAATGTGTTGTCAGAAAATCCATGCTTGATCCTGTGTA 2484  
QY 2910 TATCCAAGGGTTGAGCAGCTTCAAGCAGAGCTCCACAGTCTCAAGAGTGAATCAT 2969  
DB 2485 TATTCGTTTAAAGCTGTAATGTTTCAGAGCCCTGCACACATGTAATTTGTGTGT 2544  
QY 2970 CTTCCTACTAAAGAGATGATCGCTTGTCTAAGACCTATCTGTGGAGACTAAGACGG 3029  
DB 2545 ATTCCTCAGAAAAGAAAAGACCTCATCCGAATGATGTTTGGAGTGTATTTGATGG 2604  
QY 3030 CGATATGGCTGGGTCTGCTGGGATCCGAGATCGTC 3066  
DB 2605 GGATATCTACTTTGTTGCTGGATCAAGACATGATC 2641  
Search completed: August 11, 2005, 00:49:20  
Job time : 17526 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2005, 12:28:07 / Search time 2057 Seconds  
(without alignments)  
12104.245 Million cell updates/sec

Title: US-09-913-878a-1\_COPY\_2447\_6652

Perfect score: 4206  
Sequence: 1 atgaaccttacttactctag.....tcacaggggaatgagattat 4206

Scoring table: IDENTITY\_NUC  
Gapop 10.0, Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N Geneseq, 16Dec04: \*  
1: geneseqn1980s: \*  
2: geneseqn1990s: \*  
3: geneseqn2000s: \*  
4: geneseqn2001as: \*  
5: geneseqn2001bs: \*  
6: geneseqn2002as: \*  
7: geneseqn2002bs: \*  
8: geneseqn2003as: \*  
9: geneseqn2003bs: \*  
10: geneseqn2003cs: \*  
11: geneseqn2003ds: \*  
12: geneseqn2004as: \*  
13: geneseqn2004bs: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4206	100.0	8045	3	AAA65171 Neurospor
2	4202.8	99.9	4206	12	ADN40832 Neurospor
3	3078	73.2	3078	12	ADN40834 Truncated
4	93.2	2.2	615	8	ABZ54843 Aspergill
5	52.4	1.2	2000	8	ADA71938 Rice gene
6	51.8	1.2	2000	8	ADA71938 Rice gene
7	51.6	1.2	1164	4	AAQ03159 Complete
8	49	1.2	3731	4	AA517845 Tomato RN
9	49	1.2	3731	4	AA517837 Tomato RN
10	49	1.2	3731	4	AA517837 Tomato RN
11	45.6	1.1	495	9	ACH16195 Human adu
12	45.6	1.1	2710	4	AA844447 Nucleotid
13	44	1.0	282	6	ABL71773 Corn tass
14	44	1.0	1050	4	AA633132 DNA encod
15	44	1.0	1318	5	AA633132 DNA encod
16	44	1.0	1540	3	AA633132 DNA encod
17	44	1.0	1555	4	AA633132 DNA encod
18	44	1.0	1831	6	AA633132 DNA encod
19	44	1.0	1926	5	AA633132 DNA encod
20	44	1.0	1959	5	AA633132 DNA encod

21	44	1.0	3170	4	AA65216 Human imm
22	44	1.0	3364	4	AA65217 Human imm
23	44	1.0	3901	3	AA65217 Human imm
C 24	44	1.0	94330	11	ACN44662 Human gen
C 25	42.4	1.0	506	12	ACH87424 Human gen
26	42.4	1.0	1168	10	ACH87424 Human gen
27	42.2	1.0	594	5	AA574221 DNA encod
28	42	1.0	1167	1	AA574221 DNA encod
29	41.2	1.0	1485	4	AA554166 Pseudomon
30	41.2	1.0	1485	4	AA554166 Pseudomon
31	41.2	1.0	1535	4	AA546197 Human DNA
C 32	41.2	1.0	1535	8	ACA89647 Human enco
C 33	41.2	1.0	1535	8	ACA89647 Human enco
C 34	41.2	1.0	1535	8	ACA73657 Human enco
C 35	41.2	1.0	1535	8	ACA05972 Human sec
C 36	41.2	1.0	1535	8	ACA66806 CDNA enco
C 37	41.2	1.0	1535	8	ACF20381 Human sec
C 38	41.2	1.0	1535	8	ACF19767 Human sec
C 39	41.2	1.0	1535	8	ACD22055 Human sec
C 40	41.2	1.0	1535	8	ACF13220 Human sec
C 41	41.2	1.0	1535	8	ACD25323 Human sec
C 42	41.2	1.0	1535	8	ACF00372 Human sec
C 43	41.2	1.0	1535	8	ACA72429 Novel hum
C 44	41.2	1.0	1535	8	ACD04953 Novel hum
C 45	41.2	1.0	1535	8	ACD18414 Human sec
			1535	8	ACD08421 Human sec

## ALIGNMENTS

RESULT 1	AAA65171	standard; DNA; 8045 BP.
ID	AAA65171	
XX	AAA65171;	
AC		
XX		
DT	16-NOV-2000	(first entry)
XX		
DE	Neurospora crassa gde-1 gene.	
XX		
KW	Gene silencing; quelling deficient; gde-1; al-1; ds.	
XX		
OS	Neurospora crassa.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	2447..6655
FT		/*tag= a
FT		/product= "ODE-1"
XX		
PN	MO200050581-A2.	
XX		
PD	31-AUG-2000.	
XX		
PF	16-FEB-2000; 2000MO-IT000048.	
XX		
PR	22-FEB-1999; 99IT-RM000117.	
XX		
PA	(UTRO-) UNIV ROMA LA SAPIENZA.	
XX		
PI	Macino G, Cogoni C;	
XX		
DR	WPI; 2000-579171/54.	
DR	P-85DB; AAB13956.	
XX		
PT	Novel polynucleotide encoding a polypeptide which has a silencing	
PT	activity and comprising a RNA-dependent RNA polymerase domain.	
XX		
PS	Claim 1; Page 31-43; 48pp; English.	
XX		
CC	The present sequence is the Neurospora crassa gde-1 gene. This gene has	
CC	silencing activity. The gde-1 gene was isolated by mutational analysis of	
CC	an al-1 transgenic strain. This strain had an albino phenotype resulting	
CC	from post-transcriptional silencing of the endogenous al-1 gene.	



CC Reversion of this phenotype indicated a mutation in a silencing gene. The  
CC silencing gene, qde-1, could then be isolated. Modulation of qde-1  
CC expression may be used to inactivate genes and to silence suppression of  
CC genes

XX Sequence 8045 BP; 2089 A; 2180 C; 1950 G; 1826 T; 0 U; 0 Other;

Query Match 100.0%; Score 4206; DB 3; Length 8045;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4206; Conservaive 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 ATGAACCTTATCTCTAGAGAGAGATAGCCCGCTCGAGGAATCATTAACCGGCTC 60
Db 2447 ATGAACCTTATCTCTAGAGAGAGATAGCCCGCTCGAGGAATCATTAACCGGCTC 2506
QY 61 AATAACGACTACAACTTGCGGCTTCCAGTGTGCGACACAACTTACCCGCCACGC 120
Db 2507 AATAACGACTACAACTTGCGGCTTCCAGTGTGCGACACAACTTACCCGCCACGC 2566
QY 121 CGGAGAGAGCTGGCCGAGAGTACAGAGATTTGGGTGCCATGACAAAGATCTACAGACC 180
Db 2567 CGGAGAGAGCTGGCCGAGAGTACAGAGATTTGGGTGCCATGACAAAGATCTACAGACC 2626
QY 181 CTGAACCTTCTACTAGCGGAGAGATGACTCCCTGAACGAGCGAGAGCACTTCTTC 240
Db 2627 CTGAACCTTCTACTAGCGGAGAGATGACTCCCTGAACGAGCGAGAGCACTTCTTC 2686
QY 241 ATGAGGCGCAAGCTGCGAGCTCGAAGCTGGTCCCAAGCCGACGCGCAGC 300
Db 2687 ATGAGGCGCAAGCTGCGAGCTCGAAGCTGGTCCCAAGCCGACGCGCAGC 2746
QY 301 CTTCCTGCTGTCGAAGAACTCCCGCGCGCTACTGCGCGCAACATGGGCAATTGCG 360
Db 2747 CTTCCTGCTGTCGAAGAACTCCCGCGCGCTACTGCGCGCAACATGGGCAATTGCG 2806
QY 361 ACTGCTGCTCGAGGCTTAATAGGTTATGCACTCCCAATACAGACGAGTGA 420
Db 2807 ACTGCTGCTCGAGGCTTAATAGGTTATGCACTCCCAATACAGACGAGTGA 2866
QY 421 ACGTTTGGCAGAACTCTAAGCGGCGCAAGTGCGCTGAGCGGCCAACCTTACCAACCC 480
Db 2867 ACGTTTGGCAGAACTCTAAGCGGCGCAAGTGCGCTGAGCGGCCAACCTTACCAACCC 2926
QY 481 AAACGCGAGAGATGAGCCCGCAATGTCACTTTGGCTGATCCGCCCAAGCGTGGTGA 540
Db 2927 AAACGCGAGAGATGAGCCCGCAATGTCACTTTGGCTGATCCGCCCAAGCGTGGTGA 2986
QY 541 CGGCTGCGCAGAGTCTCTAATTTCAAGGCGCGCATACCCCTAAAGTCCCGCATCCA 600
Db 2987 CGGCTGCGCAGAGTCTCTAATTTCAAGGCGCGCATACCCCTAAAGTCCCGCATCCA 3046
QY 601 GTGAATACCGGTTCCAAACGACATCTCGAGAGTGAAGATCTCAATCAGTGACCAAG 660
Db 3047 GTGAATACCGGTTCCAAACGACATCTCGAGAGTGAAGATCTCAATCAGTGACCAAG 3106
QY 661 CGGAGCGCAAGGCGAAGCTGTCTGATATGTTGGCGCTGCGCGCGCGCGCGTGGCTAT 720
Db 3107 CGGAGCGCAAGGCGAAGCTGTCTGATATGTTGGCGCTGCGCGCGCGCGCGTGGCTAT 3166
QY 721 GCGAGCGCTTTGACAAAGATACCGACTCGAAGGATGCCAAATACAGAGATCCCAAGGCG 780
Db 3167 GCGAGCGCTTTGACAAAGATACCGACTCGAAGGATGCCAAATACAGAGATCCCAAGGCG 3226
QY 781 ACAGGTATAGACGAGCGGACAGGTGATTTCTTTGATATCATCTCAAGGCACTTCTAT 840
Db 3227 ACAGGTATAGACGAGCGGACAGGTGATTTCTTTGATATCATCTCAAGGCACTTCTAT 3286
QY 841 GGTTCAGTGTCTTACAGCGCTTGGCGGTCAACATTCAGACATACCCAGAGTAGTTTGA 900
Db 3287 GGTTCAGTGTCTTACAGCGCTTGGCGGTCAACATTCAGACATACCCAGAGTAGTTTGA 3346
QY 901 GGTTCCTCTTCAAGCGCGAGAGAGAGCGGCTGTGATGCCACGCTTTGAGGCTGA 960
Db 3347 GGTTCCTCTTCAAGCGCGAGAGAGAGCGGCTGTGATGCCACGCTTTGAGGCTGA 3406
QY 961 CACTTATGATAGTCTCTTAGCAAGAGAGAACCAACAGTCCCACTAGATTAACGAGCC 1020
Db 3407 CACTTATGATAGTCTCTTAGCAAGAGAGAACCAACAGTCCCACTAGATTAACGAGCC 3466
QY 1021 CTTTCAATGCTTCCCAAGGTTGAACCTTGTGATGACATTAATGATGCTTTCCAGT 1080
Db 3467 CTTTCAATGCTTCCCAAGGTTGAACCTTGTGATGACATTAATGATGCTTTCCAGT 3526
QY 1081 TCCGCGCGGCGAGGCGCAATTTCCGAGCGGAGTCCGCTCAAAATGATCGCTCGAGAGAA 1140
Db 3527 TCCGCGCGGCGAGGCGCAATTTCCGAGCGGAGTCCGCTCAAAATGATCGCTCGAGAGAA 3586
QY 1141 GAAAGCGCTGATCTCAGGTTCAAGTTCAATGCCCGGTGGTTGACAGCTCGGCTAGAAAT 1200
Db 3587 GAAAGCGCTGATCTCAGGTTCAAGTTCAATGCCCGGTGGTTGACAGCTCGGCTAGAAAT 3646
QY 1201 ATTTGGCCGAAATTTCCCAATGCTACACAGAGTCTCTCGCTTTGATGGAAAT 1260
Db 3647 ATTTGGCCGAAATTTCCCAATGCTACACAGAGTCTCTCGCTTTGATGGAAAT 3706
QY 1261 ACCAGACTTTTATGACTGCAAGTGAAGCTTGGAGAGAGAGGCGTGGGCTTAAAGTAC 1320
Db 3707 ACCAGACTTTTATGACTGCAAGTGAAGCTTGGAGAGAGAGGCGTGGGCTTAAAGTAC 3766
QY 1321 GACCTTCTGATCTACCGGCGCGATGTCAACAGATATGTGAAGA CTCTTACCGGCTT 1380
Db 3767 GACCTTCTGATCTACCGGCGCGATGTCAACAGATATGTGAAGA CTCTTACCGGCTT 3826
QY 1381 GATGCTTTCCGTGTGAACCTTTTCCAGAAAGCCGCCCAAGAGTGTTCGTGACCGCA 1440
Db 3827 GATGCTTTCCGTGTGAACCTTTTCCAGAAAGCCGCCCAAGAGTGTTCGTGACCGCA 3886
QY 1441 ATGAGCGGGAACCTTTAGAGCAAGGATAGCGGCTTGTCTCTGCTGTTCTAGACTAC 1500
Db 3887 ATGAGCGGGAACCTTTAGAGCAAGGATAGCGGCTTGTCTCTGCTGTTCTAGACTAC 3946
QY 1501 AATCCGAGCAACTCGGCTACTGCGCCCTTTACTCTTGTGAAGTGAAGCGCTCATGTTT 1560
Db 3947 AATCCGAGCAACTCGGCTACTGCGCCCTTTACTCTTGTGAAGTGAAGCGCTCATGTTT 4006
QY 1561 GAGCAGGCTGTGATCTACCGGCTGGTGGTCTGATAGTTTTCAGATCCTTAAT 1620
Db 4007 GAGCAGGCTGTGATCTACCGGCTGGTGGTCTGATAGTTTTCAGATCCTTAAT 4066
QY 1621 CGGTGCTTACAGAGCAAGGCGGCAAGTGAACGCGGCTGACAGCAACCAAGGTTGG 1680
Db 4067 CGGTGCTTACAGAGCAAGGCGGCAAGTGAACGCGGCTGACAGCAACCAAGGTTGG 4126
QY 1681 GTGAAGAGTCAATCCAGTGGCTCAGATGGGCAACATTTCTTTGTAGGCGGCCAATGG 1740
Db 4127 GTGAAGAGTCAATCCAGTGGCTCAGATGGGCAACATTTCTTTGTAGGCGGCCAATGG 4186
QY 1741 CGGCTTTTCTTGGCCAAAGATGCCGATACAGAAACTCTCAGGAGTTTCCAGCTCCG 1800
Db 4187 CGGCTTTTCTTGGCCAAAGATGCCGATACAGAAACTCTCAGGAGTTTCCAGCTCCG 4246
QY 1801 GCGAGAGCCCGAAACCCATCATCAAGAGAGAGTCACTTCTTTCCGAGACCGGCAAT 1860
Db 4247 GCGAGAGCCCGAAACCCATCATCAAGAGAGAGTCACTTCTTTCCGAGACCGGCAAT 4306
QY 1861 ACGTTCGACCTGATGTGTTCAAGAGAGATCTGTCTTCGCGAGAGAACTCTGAGAG 1920
Db 4307 ACGTTCGACCTGATGTGTTCAAGAGAGATCTGTCTTCGCGAGAGAACTCTGAGAG 4366
QY 1921 CAAAGGACCGAGTTCAAGATTAATGTAATCTGAGCTGCTCTGCACTGACAAACA 1980
Db 4367 CAAAGGACCGAGTTCAAGATTAATGTAATCTGAGCTGCTCTGCACTGACAAACA 4426
QY 1981 ACTTGGCAGCGGCACTCAAGTTTCTCCGATACACTGGTGTGATGAAGATAT 2040
Db 4427 ACTTGGCAGCGGCACTCAAGTTTCTCCGATACACTGGTGTGATGAAGATAT 4486
```



2041 GCCATTATGACATTGAGGCTCACCAGATCAGACACCAAGACCGATCTTTGACCT 2100  
4487 GCCATTATGACATTGAGGCTCACCAGATCAGACACCAAGACCGATCTTTGACCT 4546  
2101 TCAGGCACTGGCGAAGTAGATGAATGACGCTGTAGCCGCATGTGCGAAGCGTGGCCAG 2160  
4547 TCAGGCACTGGCGAAGTAGATGAATGACGCTGTAGCCGCATGTGCGAAGCGTGGCCAG 4606  
2161 AGGATAGCGCATGTTCTCGGTTTGGGTGATGTGCTCTGTGTGCAAGGGCGGTTGGT 2220  
4607 AGGATAGCGCATGTTCTCGGTTTGGGTGATGTGCTCTGTGTGCAAGGGCGGTTGGT 4666  
2221 TCAGGCAAGGGAATGTGGGTTATCGAGCTTGAACAGACAGGCGATGAGATTGGATCGAG 2280  
4667 TCAGGCAAGGGAATGTGGGTTATCGAGCTTGAACAGACAGGCGATGAGATTGGATCGAG 4726  
2281 ACATACCCTGCTCCAGCGCAAGTGGAAATGCGACTTGTGATMAACATCAACGTACCTC 2340  
4727 ACATACCCTGCTCCAGCGCAAGTGGAAATGCGACTTGTGATMAACATCAACGTACCTC 4786  
2341 GAAGTCGGAGCGTGGCTTCTGAACCTGAATCAGCTGTCTCAACCTACAGCTTTACT 2400  
4787 GAAGTCGGAGCGTGGCTTCTGAACCTGAATCAGCTGTCTCAACCTACAGCTTTACT 4846  
2401 GTCCTGGAAGTAGAGCAGGGAAGAGTGAAGATGGCCAGGCAATCGGTGACTT 2460  
4847 GTCCTGGAAGTAGAGCAGGGAAGAGTGAAGATGGCCAGGCAATCGGTGACTT 4906  
2461 ATCAACGATTTTGAACGACAGTTCAGGAGCAAGCAATGCTTTGAATCGCCAGTGGAA 2520  
4907 ATCAACGATTTTGAACGACAGTTCAGGAGCAAGCAATGCTTTGAATCGCCAGTGGAA 4966  
2521 TTTTGGCCAAATGGGTTTACGAGATTTATTCAGTGGCCGCACTCGAGTACGCAAGCGCT 2580  
4967 TTTTGGCCAAATGGGTTTACGAGATTTATTCAGTGGCCGCACTCGAGTACGCAAGCGCT 5026  
2581 GTGCTTTTCTTGTGGGCTACCTGACGTAAGAGGAGACATCGAACTTTGATGAC 2640  
5027 GTGCTTTTCTTGTGGGCTACCTGACGTAAGAGGAGACATCGAACTTTGATGAC 5086  
2641 AGTGGATTGATCCCAAGAACAAAGTACTTGAACAGATCGCTGGATCTTCAAAAG 2700  
5087 AGTGGATTGATCCCAAGAACAAAGTACTTGAACAGATCGCTGGATCTTCAAAAG 5146  
2701 CGGAAATGTGACAGCTTGAATCCAGCTGAACATCCGTGTGCTGATCAGCTATCAATT 2760  
5147 CGGAAATGTGACAGCTTGAATCCAGCTGAACATCCGTGTGCTGATCAGCTATCAATT 5206  
2761 TACATGATTGCGGATTTCTGGGGTGTGCTTGAAGAAAATGAGTTCAATGTGGAATTCCTC 2820  
5207 TACATGATTGCGGATTTCTGGGGTGTGCTTGAAGAAAATGAGTTCAATGTGGAATTCCTC 5266  
2821 TCAAAATTGACAGAGCAGAGAGAGTCTTTTACACTCTATCGACTGTGATGTCTCTG 2880  
5267 TCAAAATTGACAGAGCAGAGAGAGTCTTTTACACTCTATCGACTGTGATGTCTCTG 5326  
2881 GCGGGAATCCCAAGCCATTTTCCCTAGTATATCCACGCGGTTCAAGCATCTTCAAGCCA 2940  
5327 GCGGGAATCCCAAGCCATTTTCCCTAGTATATCCACGCGGTTCAAGCATCTTCAAGCCA 5386  
2941 GAGCTCAAGATCTCAAGGATGTAATCATCTTCTACTATAAGAGATGTAACCGCTTGGT 3000  
5387 GAGCTCAAGATCTCAAGGATGTAATCATCTTCTACTATAAGAGATGTAACCGCTTGGT 5446  
3001 AAGAACTATCTGTGTGAGACTACGACGCGATATGCGCTGCTGTGCGGATCCGAG 3060  
5447 AAGAACTATCTGTGTGAGACTACGACGCGATATGCGCTGCTGTGCGGATCCGAG 5506  
3061 ATGCTCATGCTTGTGTGATATGCGGAAATGCTCTGAGCCGACCTGTCTAGGTAACCTA 3120  
5507 ATGCTCATGCTTGTGTGATATGCGGAAATGCTCTGAGCCGACCTGTCTAGGTAACCTA 5566

3121 AAGAGGACAAAACGACTTTCAAACTATTATGCGCTCACACGGGACAGGCGCTACGCGGCC 3180  
5567 AAGAGGACAAAACGACTTTCAAACTATTATGCGCTCACACGGGACAGGCGCTACGCGGCC 5626  
3181 AAGAGGACGTAACATACGATATGATCCAGAGAGCTTCCATTTTCCGCTTCAGGCCAAC 3240  
5627 AAGAGGACGTAACATACGATATGATCCAGAGAGCTTCCATTTTCCGCTTCAGGCCAAC 5686  
3241 TTTCTGGGCAATGACCTAACCTAACAAAGAGGCTGTGTACATCAACATATGTTGCT 3300  
5687 TTTCTGGGCAATGACCTAACCTAACAAAGAGGCTGTGTGTACATCAACATATGTTGCT 5746  
3301 AAGAGCGGCGCATCATTTCTTACTTCACTGCTGTGGAAACCTGTGATCAAGACAAACAA 3360  
5747 AAGAGCGGCGCATCATTTCTTACTTCACTGCTGTGGAAACCTGTGATCAAGACAAACAA 5806  
3361 GGTATGTCTTTTAAACGAGACAGCTGGGCTCAATTGCTGTGGAACTGCTTGGCGGTGCA 3420  
5807 GGTATGTCTTTTAAACGAGACAGCTGGGCTCAATTGCTGTGGAACTGCTTGGCGGTGCA 5866  
3421 TTTGCTCTTCTGACCCCAATGTAACAAGGAGAGATTTGGCTCGGGGCGGGAGGCTTACC 3480  
5867 TTTGCTCTTCTGACCCCAATGTAACAAGGAGAGATTTGGCTCGGGGCGGGAGGCTTACC 5926  
3481 CACATTATGACTTACCTGAAATTTCCATGCGCAGGCTGCGATTTGACAAAGAACTGGAA 3540  
5927 CACATTATGACTTACCTGAAATTTCCATGCGCAGGCTGCGATTTGACAAAGAACTGGAA 5986  
3541 GCCTTCACAATGCCATGAAAGCGGCCAAGATACAGAGACGCGCTCATCTTTGGGAT 3600  
5987 GCCTTCACAATGCCATGAAAGCGGCCAAGATACAGAGACGCGCTCATCTTTGGGAT 6046  
3601 CCGGATCTGCTCTCTACTACGATTCAGGTTCTTCAAGGAATTTAGCGACAAAGTGGCATGCTC 3660  
6047 CCGGATCTGCTCTCTACTACGATTCAGGTTCTTCAAGGAATTTAGCGACAAAGTGGCATGCTC 6106  
3661 GCACTGCTATTACGACTCTGAAGAACCGTATCGGCGAAGTTCGAAAGAAATATGCAAG 3720  
6107 GCACTGCTATTACGACTCTGAAGAACCGTATCGGCGAAGTTCGAAAGAAATATGCAAG 6166  
3721 TTGGTCAAAAACAGAGATGAGAGACAGCAAGGACCTTACCTGTCCGCTCAACGAG 3780  
6167 TTGGTCAAAAACAGAGATGAGAGACAGCAAGGACCTTACCTGTCCGCTCAACGAG 6226  
3781 GTTTATGAATAATGTGGGCTGATCAGGCTGAGGCGATGGAACAATCGGAGCAAAATAT 3840  
6227 GTTTATGAATAATGTGGGCTGATCAGGCTGAGGCGATGGAACAATCGGAGCAAAATAT 6286  
3841 GATTCTAAGGTGATCAGGTTGCTGAGCTGTCTTCTCGCGACCGTGAATGAATACA 3900  
6287 GATTCTAAGGTGATCAGGTTGCTGAGCTGTCTTCTCGCGACCGTGAATGAATACA 6346  
3901 TGGGCAATTTGCTGAGGCTTACAGGCTTTCAAGCTGTATCTAACCAAGAGCCCAAGTTC 3960  
6347 TGGGCAATTTGCTGAGGCTTACAGGCTTTCAAGCTGTATCTAACCAAGAGCCCAAGTTC 6406  
3961 GTGTGGAGATGTGGCGGACAGAGCTGCGTACATTTAAGGCGAGATGAGAGAGAGACC 4020  
6407 GTGTGGAGATGTGGCGGACAGAGCTGCGTACATTTAAGGCGAGATGAGAGAGAGACC 6466  
4021 GGTGAAGGCGCCCGGCGTTGATGACCGGCTTCAATGATCGGAGCTTGAATGCGGATTAAG 4080  
6467 GGTGAAGGCGCCCGGCGTTGATGACCGGCTTCAATGATCGGAGCTTGAATGCGGATTAAG 6526  
4081 AAGTTTACGAAGAGATGTGGCCAGGCTGGAAGGCGATGGAATCGAGTACCTGATCCG 4140  
6527 AAGTTTACGAAGAGATGTGGCCAGGCTGGAAGGCGATGGAATCGAGTACCTGATCCG 6586  
4141 GAGGCTATGAAGTGTGGGCGGAGATGATTTTGTGATGGAATTTGGTTTCAAGGGAATGGC 4200  
6587 GAGGCTATGAAGTGTGGGCGGAGATGATTTTGTGATGGAATTTGGTTTCAAGGGAATGGC 6646  
4201 GATTAT 4206

|||||  
Db 6647 GATTAT 6652

RESULT 2  
ADN40832  
ID ADN40832 standard; DNA; 4206 BP.  
XX  
AC ADN40832;  
XX  
DT 15-JUL-2004 (first entry)  
XX  
DE Neurospora crassa QDE-1 (quelling defective) DNA SeqID 1.  
XX  
KM soluble RNA polymerase; post-transcriptional gene silencing; PTGS;  
KM microarray; RNA interference; RNA binding protein; QDE-1;  
KM quelling defective; gene; ds.  
XX  
OS Neurospora crassa.  
XX  
FH Key Location/Qualifiers  
FH CDS 1..4206  
FT /\*tag= a  
FT /product= "QDE-1 protein"  
XX  
PN MO2004035784-A2.  
XX  
PD 29-APR-2004.  
XX  
PF 17-OCT-2003; 2003MO-F1000776.  
XX  
PR 21-OCT-2002; 2002US-0419562P.  
XX  
PA (RNAL-) RNA LINE OY.  
XX  
PI Mekeyev E, Bamford D;  
XX  
DR WPI; 2004-399869/37.  
XX  
DR P-PSDB; ADN40833.  
XX  
PT Producing nucleic acid product, involves contacting with template an RNA  
PT polymerase which produces short complementary RNA copies of template,  
PT that are scattered throughout entire template, and template-length  
PT complementary RNA copies.  
XX  
PS Claim 11; SEQ ID NO 1; 95bp; English.  
XX  
CC This invention relates to a novel method for producing a soluble RNA  
CC polymerase protein. Specifically, it refers to an RNA polymerase enzyme  
CC isolated from the post-transcriptional gene silencing (PTGS) pathway that  
CC is useful for in vitro RNA synthesis kits. The present invention  
CC describes an enzyme with enhanced solubility capable of producing two  
CC types of reaction products i.e. both short and long RNA copies of the  
CC template including template length complementary RNA copies. Furthermore,  
CC it can also copy single stranded DNA templates in reactions that do not  
CC require a primer for initiation of RNA synthesis. As such, this method  
CC can also be used to produce radioactively or chemically labelled RNA  
CC probes for in situ hybridisation analyses or microarray work, to generate  
CC trigger RNA molecules to induce RNA interference effects in vivo and in  
CC vitro or to study nucleic acid secondary structure, nucleic acid-protein  
CC interactions or more preferably RNA-protein interactions. This  
CC polynucleotide sequence is Neurospora crassa DNA that encodes the RNA  
CC binding protein identified as QDE-1 (quelling defective), given in an  
CC exemplification of the invention.  
XX  
SQ Sequence 4206 BP; 1038 A; 1141 C; 1113 G; 914 T; 0 U; 0 Other;

Query Match 99.9%; Score 4202.8; DB 12; Length 4206;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4204; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 ATGAACCTTATTCTCTAGAGAGAAATGCCCCCTCGAGAAATCATTAACCGGCTC 60  
|||||

Db 1 ATGAACCTTATTCTCTAGAGAGAAATGCCCCCTCGAGAAATCATTAACCGGCTC 60  
OY  
61 AATAACGACTTACCAACTGGGCTCTCAAGTGTGCGAGACAACTCTCAACCCGCCACGG 120  
Db 61 AATAACGACTTACCAACTGGGCTCTCAAGTGTGCGAGACAACTCTCAACCCGCCACGG 120  
OY  
121 CGGAGAGAGCTGGCCGAGAGTGAAGAGATTTTCGGTCCGCATGACAAGATCTACAGAGCC 180  
Db 121 CGGAGAGAGCTGGCCGAGAGTGAAGAGATTTTCGGTCCGCATGACAAGATCTACAGAGCC 180  
OY  
181 CTGAACCTTTCTTACTGCGGAGAGATGATCTCCCTTAACCAAGGCAAGCACTTCTTC 240  
Db 181 CTGAACCTTTCTTACTGCGGAGAGATGATCTCCCTTAACCAAGGCAAGCACTTCTTC 240  
OY  
241 ATGAGAGCCAAAGCTCGAGCTGAACTGAGTGCCTCAAGCCCAAGCCGACCTGACAGC 300  
Db 241 ATGAGAGCCAAAGCTCGAGCTGAACTGAGTGCCTCAAGCCCAAGCCGACCTGACAGC 300  
OY  
301 CTTCCGTTGCTCAAGGAACCTCCGCGCGCTACTGCGGCGCAACATGGGCAATTGACG 360  
Db 301 CTTCCGTTGCTCAAGGAACCTCCGCGCGCTACTGCGGCGCAACATGGGCAATTGACG 360  
OY  
361 ACTGTGTTGCTCGAGGTGCTTAATAGGTTATGCACTCCCAATTAACACACAGGTGCA 420  
Db 361 ACTGTGTTGCTCGAGGTGCTTAATAGGTTATGCACTCCCAATTAACACACAGGTGCA 420  
OY  
421 ACCTTTGGCAGAACTTAAGCGGCGCCAGTGGCGCTGAGCGGCCCAACTCTTACCAACCC 480  
Db 421 ACCTTTGGCAGAACTTAAGCGGCGCCAGTGGCGCTGAGCGGCCCAACTCTTACCAACCC 480  
OY  
481 AAACGCAAGAGTAGAGCCCGCAATGTCACTTTGCTGATCCGCCCAACGCTGTTGACT 540  
Db 481 AAACGCAAGAGTAGAGCCCGCAATGTCACTTTGCTGATCCGCCCAACGCTGTTGACT 540  
OY  
541 CGCTTGCCACAGAGTCTCTTATTCAGCGGCGGAGATACCCCTTAAGTTCCCGGATCA 600  
Db 541 CGCTTGCCACAGAGTCTCTTATTCAGCGGCGGAGATACCCCTTAAGTTCCCGGATCA 600  
OY  
601 GTGAATACCGGTTTCCAAAGCAACCATCTCTCGAGAGTGAATTCATCATGATGACCAAG 660  
Db 601 GTGAATACCGGTTTCCAAAGCAACCATCTCTCGAGAGTGAATTCATCATGATGACCAAG 660  
OY  
661 CGGCGCAAGGGCAAGCTGTCTGATTAATGTTGCCGTGCGCGCCGCCCGCTGCTATT 720  
Db 661 CGGCGCAAGGGCAAGCTGTCTGATTAATGTTGCCGTGCGCGCCGCCCGCTGCTATT 720  
OY  
721 GCGAGGCTTTGAGCAAGGATCCGACTCGAAGGCAATGCCAATACGAGATGCCAGGGG 780  
Db 721 GCGAGGCTTTGAGCAAGGATCCGACTCGAAGGCAATGCCAATACGAGATGCCAGGGG 780  
OY  
781 ACAAGTCATATGACGAGCGGACCGAGTGAATTCCTTGTATATCTCAAGGCACTTCTAT 840  
Db 781 ACAAGTCATATGAGAGCGGACCGAGTGAATTCCTTGTATATCTCAAGGCACTTCTAT 840  
OY  
841 GGTTCAGAGTCTTTCAGCGCTTGCGCTGCAATCAAGGCACTAACCGAGTGTGTTGAG 900  
Db 841 GGTTCAGAGTCTTTCAGCGCTTGCGCTGCAATCAAGGCACTAACCGAGTGTGTTGAG 900  
OY  
901 GCTTCCTCTTTCACAGCCAGAGAGAGAGCGGCGCTGAGTGCACCGCTTTGAGGTGGA 960  
Db 901 GCTTCCTCTTTCACAGCCAGAGAGAGAGCGGCGCTGAGTGCACCGCTTTGAGGTGGA 960  
OY  
961 CACTTATTTAGTCTCTCTAGCAAAAGAGAAACCAACCACTCCACATATAGAACAGGCC 1020  
Db 961 CACTTATTTAGTCTCTCTAGCAAAAGAGAAACCAACCACTCCACATATAGAACAGGCC 1020  
OY  
1021 CTTTCATCGTCTTTCAGGAGTGAATCTGTTGAGCACTTACTATGAGTGTGTTCCAGT 1080  
Db 1021 CTTTCATCGTCTTTCAGGAGTGAATCTGTTGAGCACTTACTATGAGTGTGTTCCAGT 1080  
OY  
1081 TCCGCGCGGAGGAGGAGCAATTCGCGAGCGAGTCCGTAATGAGACTGGCTGGAGCGAA 1140  
Db 1081 TCCGCGCGGAGGAGGAGCAATTCGCGAGCGAGTCCGTAATGAGACTGGCTGGAGCGAA 1140

1141 GAAAGCGCTGATCTCAGGTTCAAGTTCATGCCCGGATGTTGAGCTCGGCTGAGAAAT 1200  
1141 GAAAGCGCTGATCTCAGGTTCAAGTTCATGCCCGGATGTTGAGCTCGGCTGAGAAAT 1200  
1201 ATTTGGCGAAATTTCCCAATGGCTACAGAACCTCTCGCTGTTGACATGGAAATT 1260  
1201 ATTTGGCGAAATTTCCCAATGGCTACAGAACCTCTCGCTGTTGACATGGAAATT 1260  
1261 ACCGACTCTTTATGACACTGCAAGTAGACTTTGGAGACGAGACCTGGGCTTAAGTAC 1320  
1261 ACCGACTCTTTATGACACTGCAAGTAGACTTTGGAGACGAGACCTGGGCTTAAGTAC 1320  
1321 GACCTTCTGCTGCTACCGCGCGAGATGTCAGATATCTGGAGAGACTCTACCGGCTT 1380  
1321 GACCTTCTGCTGCTACCGCGCGAGATGTCAGATATCTGGAGAGACTCTCTACCGGCTT 1380  
1381 GATGCTTTCCGTGTTAAACCTTTCCAGAAAAGCCGCCAACGACGTGTTGTAACGCCA 1440  
1381 GATGCTTTCCGTGTTAAACCTTTCCAGAAAAGCCGCCAACGACGTGTTGTAACGCCA 1440  
1441 ATGACGGGCAACTTTGAGAGCAAGTAGTCCGTTGTTCTCTGCTGTTCTAGACTAC 1500  
1441 ATGACGGGCAACTTTGAGAGCAAGTAGTCCGTTGTTCTCTGCTGTTCTAGACTAC 1500  
1501 AATCGGCAACTGCGCTACCGCCCTTACCTGTTGAGAGCTGGAAGCGCTGATGTC 1560  
1501 AATCGGCAACTGCGCTACCGCCCTTACCTGTTGAGAGCTGGAAGCGCTGATGTC 1560  
1561 GAGCAGGCTGTGCACTCACCCGTCGTTCCGTCCTGATAGTTTTCGAGACTTATA 1620  
1561 GAGCAGGCTGTGCACTCACCCGTCGTTCCGTCCTGATAGTTTTCGAGACTTATA 1620  
1621 CCGTCGCTTACGAGACCAAGCCCAAGTGTACCGCGGTGTGACAAACCAACAGTGTG 1680  
1621 CCGTCGCTTACGAGACCAAGCCCAAGTGTACCGCGGTGTGACAAACCAACAGTGTG 1680  
1681 GTGCAAGAGTCATCCAGTGTGCTCAGATGGGGCAACCTCTGTTGAGGCGGCAATGG 1740  
1681 GTGCAAGAGTCATCCAGTGTGCTCAGATGGGGCAACCTCTGTTGAGGCGGCAATGG 1740  
1741 CGGCTTTCTTCCGCAAAAGATGCGGATACAGAAACCTCTCAGGAGTTCCAGCTCCG 1800  
1741 CGGCTTTCTTCCGCAAAAGATGCGGATACAGAAACCTCTCAGGAGTTCCAGCTCCG 1800  
1801 GCCGAGACCCGAAACCCATCATCAAGAGAGAGTCCACTTCTTTGCCGAGACCGGCA 1860  
1801 GCCGAGACCCGAAACCCATCATCAAGAGAGAGTCCACTTCTTTGCCGAGACCGGCA 1860  
1861 ACCTTCGCACTGATGTGTTCAGACGAGATCTGTGCTTCCGCGAGAGAACTGTAGAG 1920  
1861 ACCTTCGCACTGATGTGTTCAGACGAGATCTGTGCTTCCGCGAGAGAACTGTAGAG 1920  
1921 CAAGCAGACCGAGTTCAAGTGTAGTCAAAATGCTGAGCTGCTCTGCACTGCAACAC 1980  
1921 CAAGCAGACCGAGTTCAAGTGTAGTCAAAATGCTGAGCTGCTCTGCACTGCAACAC 1980  
1981 ACTTGGCAGCGCACTCAAGTGTGTCTCCGATACAGCTCGGCTGAGTAAACATAT 2040  
1981 ACTTGGCAGCGCACTCAAGTGTGTCTCCGATACAGCTCGGCTGAGTAAACATAT 2040  
2041 GCCATTATGACATTTGAGACCTCAACGATCAGACCAACAGACGATCTTTTCACTT 2100  
2041 GCCATTATGACATTTGAGACCTCAACGATCAGACCAACAGACGATCTTTTCACTT 2100  
2101 TCAGGCACTGGCGAAGTATGATACCGGTGTAGGCGCATGTCCGAAAGCTGGCCAG 2160  
2101 TCAGGCACTGGCGAAGTATGATACCGGTGTAGGCGCATGTCCGAAAGCTGGCCAG 2160  
2161 AGGATACCGCATGTTCTCGGTTTGGGTATGTGCTCTGCTGTGCAAGGCGGTTTGT 2220  
2161 AGGATACCGCATGTTCTCGGTTTGGGTATGTGCTCTGCTGTGCAAGGCGGTTTGT 2220  
2161 AGGATACCGCATGTTCTCGGTTTGGGTATGTGCTCTGCTGTGCAAGGCGGTTTGT 2220

2221 TCGCCAGGGAATGTGGTTATCGACTTGAACGACACAGCGGATGAGATTGATGAG 2280  
2221 TCGCCAGGGAATGTGGTTATCGACTTGAACGACACAGCGGATGAGATTGATGAG 2280  
2281 ACATACCCGTCAGCGCAAGTGGAAATGCACTTGTGTATTAACATCAAGTACCTC 2340  
2281 ACATACCCGTCAGCGCAAGTGGAAATGCACTTGTGTATTAACATCAAGTACCTC 2340  
2341 GAAGTCGGAGCGTGTCTGAACTGAAGTCAAGCTGTCTCAACCTACAGCTTTACT 2400  
2341 GAAGTCGGAGCGTGTCTGAACTGAAGTCAAGCTGTCTCAACCTACAGCTTTACT 2400  
2401 GTCCTGAAATAGAGCAGGAGCAAGAGTGAATGCGCAGGCAATCGGTGAC 2460  
2401 GTCCTGAAATAGAGCAGGAGCAAGAGTGAATGCGCAGGCAATCGGTGAC 2460  
2461 ATCAACGATTTGCAACGACAGTTCAAGCAGCAAAAGCATCTTTGAATCCGCCAGTGG 2520  
2461 ATCAACGATTTGCAACGACAGTTCAAGCAGCAAAAGCATCTTTGAATCCGCCAGTGG 2520  
2521 TTTCCGCAATGGGTTTACGAGAGTTATTCAGTCCGCACTGAGTCAGCAGCGCGT 2580  
2521 TTTCCGCAATGGGTTTACGAGAGTTATTCAGTCCGCACTGAGTCAGCAGCGCGT 2580  
2581 GTGCTTTCTTCTGCTGAGCTACCTGACAGTCAAGAGAGACACTGAATCTTTGATGAC 2640  
2581 GTGCTTTCTTCTGCTGAGCTACCTGACAGTCAAGAGAGACACTGAATCTTTGATGAC 2640  
2641 AGTGGTTGATTCGCAAGAACAAAGTACTTGCAAGACATCGCTGGATCTTCAAAAG 2700  
2641 AGTGGTTGATTCGCAAGAACAAAGTACTTGCAAGACATCGCTGGATCTTCAAAAG 2700  
2701 CGGAAATGTGACAGTTGAAGTCAAGCTGAACATCCGTCGCTGATAGATATGAT 2760  
2701 CGGAAATGTGACAGTTGAAGTCAAGCTGAACATCCGTCGCTGATAGATATGAT 2760  
2761 TACATGATTCGCAATTTCTGGGGTGCTTGAAGAAATGAGTTCAATGTCGATCTCC 2820  
2761 TACATGATTCGCAATTTCTGGGGTGCTTGAAGAAATGAGTTCAATGTCGATCTCC 2820  
2821 TCAAGTTCAAGACGAGAGAGAGTCTTTTACACTCTTATCGAGTGTGATCTCGTG 2880  
2821 TCAAGTTCAAGACGAGAGAGAGTCTTTTACACTCTTATCGAGTGTGATCTCGTG 2880  
2881 GCGGATTCGCCAGCCATTTCCCTAGATATCCAGGGTTGAGAGCTTCAAGCCA 2940  
2881 GCGGATTCGCCAGCCATTTCCCTAGATATCCAGGGTTGAGAGCTTCAAGCCA 2940  
2941 GAGCTCAGAGTCTCAAGATGTATCATCTTCTACTAAGAGATGTAACGCTTGTCT 3000  
2941 GAGCTCAGAGTCTCAAGATGTATCATCTTCTACTAAGAGATGTAACGCTTGTCT 3000  
3001 AAGAGCTATCTGTGAGAGCTACGACGCGGATATGAGCTGTGCTGGGATCCGAG 3060  
3001 AAGAGCTATCTGTGAGAGCTACGACGCGGATATGAGCTGTGCTGGGATCCGAG 3060  
3061 AAGAGCTATCTGTGAGAGCTACGACGCGGATATGAGCTGTGCTGGGATCCGAG 3120  
3061 AAGAGCTATCTGTGAGAGCTACGACGCGGATATGAGCTGTGCTGGGATCCGAG 3120  
3121 AAGAGGACAAACGACTTTCAACAACTATGAGCTCAGACGAGGCTCAGCGGCT 3180  
3121 AAGAGGACAAACGACTTTCAACAACTATGAGCTCAGACGAGGCTCAGCGGCT 3180  
3181 AAGAGGACAAACGACTTTCAACAACTATGAGCTCAGACGAGGCTCAGCGGCT 3240  
3181 AAGAGGACAAACGACTTTCAACAACTATGAGCTCAGACGAGGCTCAGCGGCT 3240  
3241 TTTCTGGCATGTGCACTTAACAAAGAAAGGTCTGTATTAACATCAATAGTGTCT 3300  
3241 TTTCTGGCATGTGCACTTAACAAAGAAAGGTCTGTATTAACATCAATAGTGTCT 3300  
3301 AACAAAGCGGATCATTTCTAGTTCACTGTTGGGAAACCTGTCATCAGAGCAAGCAA 3360

```

Dh 3301 AACAGCGGCGCATCTTCTTACCTGCTGAGAGAAACCA 3360
Qy 3361 GGTATGCTCTTTTACGAGCAAGCTGGGCTCAATTGGTAGGAACTGTTGGCGGTGCA 3420
Dh 3361 GGTATGCTCTTTTACGAGCAAGCTGGGCTCAANTGGGTAGGAACTGTTGGCGGTGCA 3420
Qy 3421 TTGTCCCTCTCTGACCAATGTAACAGAGCAAGTGGCTCGGCGCGGAGAGCTTACC 3480
Dh 3421 TTGTCCCTCTCTGACCAATGTAACAGAGCAAGTGGCTCGGCGCGGAGAGCTTACC 3480
Qy 3481 CACATTTATGACTACCTGAAATCTCATGCGCAGAGCTCGATGTAACAGAACTGGAA 3540
Dh 3481 CACATTTATGACTACCTGAAATCTCATGCGCAGAGCTCGATGTAACAGAACTGGAA 3540
Qy 3541 GCCTTCACAAATGCAATGAAAGCGCCAAAGATACAGAAACGCGCTCACTTTGGGAT 3600
Dh 3541 GCCTTCACAAATGCAATGAAAGCGCCAAAGATACAGAAACGCGCTCACTTTGGGAT 3600
Qy 3601 CCGGATCTCGCTTCTTAATAAGCTTTTCAAGAGATTAGCAAAAGTGGCATGCTCC 3660
Dh 3601 CCGGATCTCGCTTCTTAATAAGCTTTTCAAGAGATTAGCAAAAGTGGCATGCTCC 3660
Qy 3661 GCACTGCTATTCAGACTCTGAGAGAACCTGATCGGCGCAATTCAGAAAGAAATATGGAGG 3720
Dh 3661 GCACTGCTATTCAGACTCTGAGAGAACCTGATCGGCGCAATTCAGAAAGAAATATGGAGG 3720
Qy 3721 TTGGTCAAAAACAGAGATGAGAGACAGCAAGAACCTTACCCTGCGCCTCAACAG 3780
Dh 3721 TTGGTCAAAAACAGAGATGAGAGACAGCAAGAACCTTACCCTGCGCCTCAACAG 3780
Qy 3781 GTTTATGAAAAATGCTGCCATCACCGCTGAGGCGATGGAACAATCCGAGCAAAATTAT 3840
Dh 3781 GTTTATGAAAAATGCTGCCATCACCGCTGAGGCGATGGAACAATCCGAGCAAAATTAT 3840
Qy 3841 GATTCTAAGGTGATCAGGTTGCTGAGAGCTGCTCCTCGCGGACCGGTGAGATGAAATACA 3900
Dh 3841 GATTCTAAGGTGATCAGGTTGCTGAGAGCTGCTCCTCGCGGACCGGTGAGATGAAATACA 3900
Qy 3901 TGGGCAATGCTGAGGCTGACAGCGCTTCAAGCTGTAACCAAGAACCCCAAGTTTC 3960
Dh 3901 TGGGCAATGCTGAGGCTGACAGCGCTTCAAGCTGTAACCAAGAACCCCAAGTTTC 3960
Qy 3961 GTGTGGCAGATGCGGGCAGACAGCTTGCCTAATTAGGCGGAGATGACAGCAAGCCC 4020
Dh 3961 GTGTGGCAGATGCGGGCAGACAGCTTGCCTAATTAGGCGGAGATGACAGCAAGCCC 4020
Qy 4021 GGTGAAGCGCGCCCGGCTGATGACCGGCTGATGATGCGGCTTGAATGCGCGGATAG 4080
Dh 4021 GGTGAAGCGCGCCCGGCTGATGACCGGCTGATGATGCGGCTTGAATGCGCGGATAG 4080
Qy 4081 AAGTTTACGAAAGATATGTGCGCAGGCTGAGAGGCGATGATGCGAGTACCTGATCCG 4140
Dh 4081 AAGTTTACGAAAGATATGTGCGCAGGCTGAGAGGCGATGATGCGAGTACCTGATCCG 4140
Qy 4141 GAGGCTATGAAAGTGTGCGGCGATGATGATGATGATGATGATGATGATGATGATGATG 4200
Dh 4141 GAGGCTATGAAAGTGTGCGGCGATGATGATGATGATGATGATGATGATGATGATGATG 4200
Qy 4201 GATTAT 4206
Dh 4201 GATTAT 4206

```

```

XX KW soluble RNA polymerase; post-transcriptional gene silencing; PTGS;
KM microarray; RNA interference; RNA binding protein; QDE-1;
XX KM quelling defective; gene; ds.
OS Neurospora crassa.
XX
FH Key Location/Qualifiers
FT CDS 1..3078
FT /tag= a
FT /partial
FT /product= "truncated QDE-1 protein"
FT /note= "Start and stop codons are absent"
XX
PN MO2004035784-A2.
XX
XX 29-APR-2004.
XX
XX 17-OCT-2003; 2003MO-FI000776.
XX
XX 21-OCT-2002; 2002US-0419562P.
XX
XX (RNAL-) RNA LINE OY.
XX
XX Mekeyev E, Bamford D;
XX
XX WPI; 2004-399869/37.
XX
XX P-PSDB; ADN40835.
XX
XX
XX Producing nucleic acid product, involves contacting with template an RNA
PT polymerase which produces short complementary RNA copies of template,
PT that are scattered throughout entire template, and template-length
PT complementary RNA copies.
XX
XX
PS Claim 11; SEQ ID NO 3; 95pp; English.
XX
XX This invention relates to a novel method for producing a soluble RNA
CC polymerase protein. Specifically, it refers to an RNA polymerase enzyme
CC isolated from the post-transcriptional gene silencing (PTGS) pathway that
CC is useful for in vitro RNA synthesis kits. The present invention
CC describes an enzyme with enhanced solubility capable of producing two
CC types of reaction products i.e. both short and long RNA copies of the
CC template including template length complementary RNA copies. Furthermore,
CC it can also copy single stranded DNA templates in reactions that do not
CC require a primer for initiation of RNA synthesis. As such, this method
CC can also be used to produce radioactively or chemically labelled RNA
CC probes for in situ hybridisation analyses or microarray work. to generate
CC trigger RNA molecules to induce RNA interference effects in vivo and in
CC vitro or to study nucleic acid secondary structure, nucleic acid-protein
CC interactions or more preferably RNA-protein interactions. This
CC polynucleotide is a truncated Neurospora crassa DNA sequence that encodes
CC the RNA binding protein identified as QDE-1 (quelling defective), given
CC in an exemplification of the invention.
XX
SQ Sequence 3078 BP; 763 A; 785 C; 835 G; 695 T; 0 U; 0 Other;
XX
XX
XX Query Match 73.2%; Score 3078; DB 12; Length 3078;
XX Best Local Similarity 100.0%; Pred. No. 0;
XX Matches 3078; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
Qy 1129 GCTGGAGCGAAGAAAGCGCTGATCTCAGTTCAAGTTCAATGCCCGGTGGTGCAGCT 1188
Dh 1 GCTCGGAGCGAAGAAAGCGCTGATCTCAGTTCAAGTTCAATGCCCGGTGGTGCAGCT 60
Qy 1189 CGGCTGAGAAATATTTGGCGGAATTTCCCAATGCTACAGAAAGCTCTTCGTGTT 1248
Dh 61 CGGCTGAGAAATATTTGGCGGAATTTCCCAATGCTACAGAAAGCTCTTCGTGTT 120
Qy 1249 GCATGGAGATTACAGACTCTTTATGACTGCAAGTGAAGTGAAGTGAAGAGAGAGCTG 1308
Dh 121 GCATGGAGATTACAGACTCTTTATGACTGCAAGTGAAGTGAAGTGAAGAGAGAGCTG 180
Qy 1309 GGCCTAAAGTACGACCTTCTGTGTTACCGGCGCGATGTGACAGATATCTGGAAGACT 1368

```

Db 181 GGCCTAAGTACGACCTCTCTGCTACACGCGGCGATGTCACAGATATCGAAGACT 240  
Qy 1369 CTCTACCGGCTTGATGCTTTCCTGTTAAACCTTTTCAGAAAAAGCCGCCAAGACGTG 1428  
Db 241 CTCTACCGGCTTGATGCTTTCCTGTTAAACCTTTTCAGAAAAAGCCGCCAAGACGTG 300  
Qy 1429 TTGCTGACCGCAATGAGGGGCACTTTGAGAGCAAAAGTAGTGGCGTTTCTCTGCT 1488  
Db 301 TTGCTGACCGCAATGAGGGGCACTTTGAGAGCAAAAGTAGTGGCGTTTCTCTGCT 360  
Qy 1489 GTTCTAGACTACAAATCCGCAACATCGCTTACGCGCCCTTTTACTTGTGAACTGAAG 1548  
Db 361 GTTCTAGACTACAAATCCGCAACATCGCTTACGCGCCCTTTTACTTGTGAACTGAAG 420  
Qy 1549 CCGCTCATGTTTCGACAGAGGCTGTGACTACCCGTCGCTTCCGTTCTGATAGTTTTC 1608  
Db 421 CCGCTCATGTTTCGACAGAGGCTGTGACTACCCGTCGCTTCCGTTCTGATAGTTTTC 480  
Qy 1609 GAGATCCTTATACGCTGGCCTACGAGACCAAGCCCAAGTGTACCGCGGTGGTCAGGAAA 1668  
Db 481 GAGATCCTTATACGCTGGCCTACGAGACCAAGCCCAAGTGTACCGCGGTGGTCAGGAAA 540  
Qy 1669 CAACCAAGTGGCTGCAAGAGATCATCGAGCTCAAGATGGGGCAACATTTCTTGTA 1728  
Db 541 CAACCAAGTGGCTGCAAGAGATCATCGAGCTCAAGATGGGGCAACATTTCTTGTA 600  
Qy 1729 GCGCGCAATGGCGCGCTTTCTTGCCAAAGATCCCGATACAGAAACTCTCAGAGAG 1788  
Db 601 GCGCGCAATGGCGCGCTTTCTTGCCAAAGATCCCGATACAGAAACTCTCAGAGAG 660  
Qy 1789 TTCCAGCTCCGCGCGAGAACCCGAAACCATTCATCAAGAGAGATCCCTTTTGCC 1848  
Db 661 TTCCAGCTCCGCGCGAGAACCCGAAACCATTCATCAAGAGAGATCCCTTTTGCC 720  
Qy 1849 GAGACCGGCAATTAAGTTCCGACTGATGTGTTCAAGAAGAGATGTGTTCCGCGAGAG 1908  
Db 721 GAGACCGGCAATTAAGTTCCGACTGATGTGTTCAAGAAGAGATGTGTTCCGCGAGAG 780  
Qy 1909 GAACCTGTAGAGCAACGAGCCGAGTTCAAAAGTAGTCAAAATGCTGAGCTGCTTCCAA 1968  
Db 781 GAACCTGTAGAGCAACGAGCCGAGTTCAAAAGTAGTCAAAATGCTGAGCTGCTTCCAA 840  
Qy 1969 CTGCAACAACAACCTTGGCAGCGGCACTCAAGTTTCTCCGATCCAGCTGGGTCTG 2028  
Db 841 CTGCAACAACAACCTTGGCAGCGGCACTCAAGTTTCTCCGATCCAGCTGGGTCTG 900  
Qy 2029 AGTAAACATATGCAATTTATGACATTTGAGGCTCAGCAGATCAGACACACAAGCCGAT 2088  
Db 901 AGTAAACATATGCAATTTATGACATTTGAGGCTCAGCAGATCAGACACACAAGCCGAT 960  
Qy 2089 CTTCTTTTCACTTCAAGGCACTGCGGAAGTGAATGAATGACGCTGAGCGCGCATCTCGGA 2148  
Db 961 CTTCTTTTCACTTCAAGGCACTGCGGAAGTGAATGAATGACGCTGAGCGCGCATCTCGGA 1020  
Qy 2149 AGCGTGGCCAAAGAGATACGAGATGTTCTCGTTTGGTGAATGGCCCTGCTGCTGAA 2208  
Db 1021 AGCGTGGCCAAAGAGATACGAGATGTTCTCGTTTGGTGAATGGCCCTGCTGCTGAA 1080  
Qy 2209 GGGCGGTTTGGTTGGCCAAAGGAATGAGTTATTCGACTTTGACGACACAGCGGATGAG 2268  
Db 1081 GGGCGGTTTGGTTGGCCAAAGGAATGAGTTATTCGACTTTGACGACACAGCGGATGAG 1140  
Qy 2269 GATTGATCGAGACATACCGCTCCACGCGCAAGTGGGAATCGACTTGTGTTGATTAACAT 2328  
Db 1141 GATTGATCGAGACATACCGCTCCACGCGCAAGTGGGAATCGACTTGTGTTGATTAACAT 1200  
Qy 2329 CAAGTACCTTCGAAGTCCGAGGCGGTCTCTGAACGATGAGCTGCTGATTAACCTA 2388  
Db 1201 CAAGTACCTTCGAAGTCCGAGGCGGTCTCTGAACGATGAGCTGCTGATTAACCTA 1260  
Qy 2389 CAGCTTTACTGTCTGGAAGATAGAGCCAGGACAAAGTGAAGATGCGCCAGGCAATC 2448

Db 1261 CAGCTTTACTGTCTGGAAGATAGAGCCAGGGAACAAGTGAAGATGCCACAGCAATC 1320  
Qy 2449 GGTGACCGTCTTATCAACGATTTTGCAACGACAGTTCAAGCGACAAAAGCATGTCTTGAAT 2508  
Db 1321 GGTGACCGTCTTATCAACGATTTTGCAACGACAGTTCAAGCGACAAAAGCATGTCTTGAAT 1380  
Qy 2509 CGCCAGTGGAAATTTGCGCAATGGGTTTACGAGAGTTATTCAGATCGCGCACTCGAGTC 2568  
Db 1381 CGCCAGTGGAAATTTGCGCAATGGGTTTACGAGAGTTATTCAGATCGCGCACTCGAGTC 1440  
Qy 2569 AGCCACGGCCGTGTGCTTTTCTTGCGGCTACCTGACAGTCAAGAGAGACATCGAAC 2628  
Db 1441 AGCCACGGCCGTGTGCTTTTCTTGCGGCTACCTGACAGTCAAGAGAGACATCGAAC 1500  
Qy 2629 TTCTGTATGAACAGTGGGTTCCATCCCAAGAACAAAGTATTTTGCAAGCATCGCCTGG 2688  
Db 1501 TTCTGTATGAACAGTGGGTTCCATCCCAAGAACAAAGTATTTTGCAAGCATCGCCTGG 1560  
Qy 2689 GATCTTCAAAAGCGAAATGTGACACGTTGAAGTCCAAAGCTGAACATCCGTGCGGTGCA 2748  
Db 1561 GATCTTCAAAAGCGAAATGTGACACGTTGAAGTCCAAAGCTGAACATCCGTGCGGTGCA 1620  
Qy 2749 TCAGCATACATTTACATGATTCGCGATTTCTGGGGTGTGCTTGAAGAAATGAGTTTCT 2808  
Db 1621 TCAGCATACATTTACATGATTCGCGATTTCTGGGGTGTGCTTGAAGAAATGAGTTTCT 1680  
Qy 2809 GTGCGATTTCTCTCAAAAGTTCAAGACAGAGAGAGCTTTTACATCTCTATCGGACTGT 2868  
Db 1681 GTGCGATTTCTCTCAAAAGTTCAAGACAGAGAGAGCTTTTACATCTCTATCGGACTGT 1740  
Qy 2869 GATGCTCTGCTGGCGCGATCCCGAGCCCAATTCCTCTAGTATATCCAAAGGTTTCAGCA 2928  
Db 1741 GATGCTCTGCTGGCGCGATCCCGAGCCCAATTCCTCTAGTATATCCAAAGGTTTCAGCA 1800  
Qy 2929 GTCTTCAAGCCAGAGCTCCACAGTCTCAAGATGTATCATCTTCTACTTAAAGAGAT 2988  
Db 1801 GTCTTCAAGCCAGAGCTCCACAGTCTCAAGATGTATCATCTTCTACTTAAAGAGAT 1860  
Qy 2989 GTACCGTTTGTCTAAGAGCTATCTGTGTGAGACTACGACGGCGATATGGGCTGGGTGTC 3048  
Db 1861 GTACCGTTTGTCTAAGAGCTATCTGTGTGAGACTACGACGGCGATATGGGCTGGGTGTC 1920  
Qy 3049 TGGGATCCGAGATTCGTCGATGTTTCGTCATATGCGGAAATGCTCGAGCCGACCTG 3108  
Db 1921 TGGGATCCGAGATTCGTCGATGTTTCGTCATATGCGGAAATGCTCGAGCCGACCTG 1980  
Qy 3109 TCTAGTACCTTAAAGAGAGCAAAACGACTTTCAACCACTTATGCTTCAACGAGCAG 3168  
Db 1981 TCTAGTACCTTAAAGAGAGCAAAACGACTTTCAACCACTTATGCTTCAACGAGCAG 2040  
Qy 3169 GGTCTAGCGGCCAAAGAGAGACTACATACGATATGATCCAGAAAGCTTCCATTTGCGC 3228  
Db 2041 GGTCTAGCGGCCAAAGAGAGACTACATACGATATGATCCAGAAAGCTTCCATTTGCGC 2100  
Qy 3229 CTGAGGCCCAACTTCTTGGGCAATGTGCACTTAACTTAAAGAAAGGCTCTGTATCATCAAC 3288  
Db 2101 CTGAGGCCCAACTTCTTGGGCAATGTGCACTTAACTTAAAGAAAGGCTCTGTATCATCAAC 2160  
Qy 3289 AATAGTGTCTTAAACAAGCCGACATCATTTCTTATGTTCACTGTGTGGGAAACCTGTGAT 3348  
Db 2161 AATAGTGTCTTAAACAAGCCGACATCATTTCTTATGTTCACTGTGTGGGAAACCTGTGAT 2220  
Qy 3349 CAGAGCAAGCAAGTATGTCTTTTAAAGCAAGCAAGCTGGGCTCAATTTGCTTAAAGGAGCTG 3408  
Db 2221 CAGAGCAAGCAAGTATGTCTTTTAAAGCAAGCAAGCTGGGCTCAATTTGCTTAAAGGAGCTG 2280  
Qy 3409 CTTGGCGGTGATTTGCTTCTCTGACCCCAATGTACAAAGAGCAAGTGGCTCGGCGC 3468  
Db 2281 CTTGGCGGTGATTTGCTTCTCTGACCCCAATGTATCAAGAGCAAGTGGCTCGGCGC 2340  
Qy 3469 GAGAGGCTTACCAACATTAATGACTACCTGAAAATCTTCATGCGCCAGGCTGCGATTTGAC 3528  
Db 2341 GAGAGGCTTACCAACATTAATGACTACCTGAAAATCTTCATGCGCCAGGCTGCGATTTGAC 2400

```

QY 3529 AAGGAATGGAAGCTTCCAAATGCCATGAAACGGCCCAAGATACAGAAACGCGCT 3588
D 2401 AAGGAATGGAAGCTTCCAAATGCCATGAAACGGCCCAAGATACAGAAACGCGCT 2460
QY 3589 CACTTTGGGATCCGGATCTCGCTCTCTACTACAGCTTTTCAAGAGATTAAGGACAG 3648
D 2461 CACTTTGGGATCCGGATCTCGCTCTCTACTACAGCTTTTCAAGAGATTAAGGACAG 2520
QY 3649 TCGGATCGTCCGATGCTATTCAGACTCTGAAAGACCGTATCCGGAAAGTCAGAAA 3708
D 2521 TCGGATCGTCCGATGCTATTCAGACTCTGAAAGACCGTATCCGGAAAGTCAGAAA 2580
QY 3709 GAATATGGCAGGTTGTCAAAAACAAGAGATGAGACAGCAAGAACCCCTTACCTGTG 3768
D 2581 GAATATGGCAGGTTGTCAAAAACAAGAGATGAGACAGCAAGAACCCCTTACCTGTG 2640
QY 3769 CGGCTCAACCAAGTTTATGAAAAATGTCGCCATCAAGCTTGAAGCGATGACAAATCC 3828
D 2641 CGGCTCAACCAAGTTTATGAAAAATGTCGCCATCAAGCTTGAAGCGATGACAAATCC 2700
QY 3829 GGAGCAATATGATCTTAAGGTGATCAGGTTGCTGAGGCTGTCCTTCCGCGGACCGT 3888
D 2701 GGAGCAATATGATCTTAAGGTGATCAGGTTGCTGAGGCTGTCCTTCCGCGGACCGT 2760
QY 3889 GAGATGAATACATGGGCTTCTGAGGGCTAGCAGCGCTTTCAGAGCTGTACTACCAACAG 3948
D 2761 GAGATGAATACATGGGCTTCTGAGGGCTAGCAGCGCTTTCAGAGCTGTACTACCAACAG 2820
QY 3949 AGCCCCCAAGTTCTGTGTGGCGATGCGCGGACAGACGCTCCGCTACATTAAGCGCGATG 4008
D 2821 AGCCCCCAAGTTCTGTGTGGCGATGCGCGGACAGACGCTCCGCTACATTAAGCGCGATG 2880
QY 4009 ACAGACGACCCCGGTGAAGCGCCCGCGCTTGAATGACCGCGTTCAATGATGCGGCTTG 4068
D 2881 ACAGACGACCCCGGTGAAGCGCCCGCGCTTGAATGACCGCGTTCAATGATGCGGCTTG 2940
QY 4069 ATGCGCGATAGAAGTTTACGAAGCAGTATGTGGCCAGGCTGAGGGCGATGATCGAG 4128
D 2941 ATGCGCGATAGAAGTTTACGAAGCAGTATGTGGCCAGGCTGAGGGCGATGATCGAG 3000
QY 4129 TACCTGTATCCGAGAGCTTATGAAAGTCTGGCGGATGATGATTTTGAATTTGTTTC 4188
D 3001 TACCTGTATCCGAGAGCTTATGAAAGTCTGGCGGATGATGATTTTGAATTTGTTTC 3060
QY 4189 ACAGGAATGCGGATAT 4206
D 3061 ACAGGAATGCGGATAT 3078

```

## RESULT 4

```

ABZ54843
ID ABZ54843 standard; cDNA; 615 BP.
XX
AC ABZ54843;
XX
DT 28-MAR-2003 (first entry)
XX
DE Aspergillus oryzae polynucleotide seq ID NO 3956.
XX
KW Aspergillus oryzae; fermentation; fungus; industrial; EST;
XX expressed sequence tag; gene; ss.
XX
OS Aspergillus oryzae.
XX
PN W0200279476-A1.
XX
PD 10-OCT-2002.
XX
PF 22-MAR-2002; 2002MO-IB000890.
XX
PR 30-MAR-2001; 2001JP-00098371.
XX

```

```

PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA (NARE-) NAT RES INST BREWING.
PA (NORO) NAT FOOD RES INST MIN AGRIC.
XX
PI Machida M, Akita O, Kashiwagi Y, Kitamoto K, Horiuchi H;
PI Takeuchi M, Kobayashi T, Kitamoto N, Gomi K, Abe K;
XX
DR WPI; 2003-046817/04.
XX
PT Detection of expression of specific Aspergillus genes for monitoring the
PT fermentation and growth conditions of the fungus, using DNA probes.
XX
PS Claim 1; SEQ ID NO 3956; 48bp + Sequence Listing; Japanese.
XX
CC The invention relates to a polynucleotide having any of 6006 specific
CC sequences (ABZ50888-ABZ56893), which are expressed by a fungus under
CC specific culture conditions including one or more of eutrophic,
CC oligotrophic, solid, early germination, alkaline, high temperature, low
CC temperature or maltose culture or polynucleotides stringently hybridizing
CC to these sequences. The polynucleotides are useful for monitoring the
CC progress of fermentation and the growth conditions of a fungus,
CC especially of Aspergillus oryzae which is widely used in industrial
CC fermentation. Also monitoring for fungal contamination. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pot_sequences
XX
SQ Sequence 615 BP; 151 A; 148 C; 162 G; 154 T; 0 U; 0 Other;

```

```

Query Match 2.2%; Score 93.2; DB 8; Length 615;
Best Local Similarity 64.2%; Pred. No. 3.5e-17;
Matches 140; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

```

```

QY 2869 GATGTCCTGTGGCGGATCCCGACCCCATTTCCCTAGATATCAACGGGTTGAGCA 2928
D 74 GACGTTCTTTTACTCTGCTCCCGCAGATGTGCTCTGATATCCAGGACAGAGCT 133
QY 2929 GTCTTCAAGCAGAGCTCCACAGTCTCAAGAGATGTAATCATCTTCTACTAAGAGAT 2988
D 134 GTGTGAGACGACAGCTCCGGGATTTCAAAAGATGTAATGTTTCCCTACTGTGTGGACA 193
QY 2989 GTTCCGCTTGTCTAAGAGCTATGTGTGAGACTAGACGGCGATGTGGCTGGCTTCG 3048
D 194 ACGCCACATCGCGCAGCATGTGTAATCCGCTGTGTGATTAAGACGGGTGATCAACGTGGGTTTC 253
QY 3049 TGGGATCCGAGATCGTCAATGTTTCTGCAATGCGGA 3086
D 254 TGGGATCCGAGATCGTCAATGTTTCTGCAATGCGGA 291

```

## RESULT 5

```

ADA71938
ID ADA71938 standard; DNA; 2000 BP.
XX
AC ADA71938;
XX
DT 20-NOV-2003 (first entry)
XX
DE Rice gene, SEQ ID 5263.
XX
KW Plant; bacterial infection; fungal infection; viral infection; rice;
XX gene; ds.
XX
OS Oryza sativa.
XX
PN W02003000898-A1.
XX
PD 03-JAN-2003.
XX
PF 22-JUN-2001; 2001MO-IB001105.
XX
PR 22-JUN-2001; 2001MO-IB001105.
XX

```



(SYGN ) SYNGENTA PARTICIPATIONS AG.

PA Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX WPI; 2003-175290/17.

DR Identifying at least one gene involved in plant resistance or response to  
PT pathogenic infection for conferring resistance or tolerance to a plant to  
PT bacterial, fungal or viral infection by determining or detecting plant  
PT gene expression.

PS Claim 27; SEQ ID NO 5263; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes  
CC involved in plant resistance or response to pathogenic infection. M1  
CC comprises identifying a gene whose expression is significantly altered in  
CC the incompatible interaction of plant gene expression relative to  
CC expression of the gene in an uninfected plant, in a mutant plant that  
CC does not express a gene associated with response to pathogenic infection,  
CC or in a corresponding incompatible or compatible interaction. (M1) is  
CC useful for conferring resistance to resistance or tolerance to a plant to  
CC bacterial, fungal or viral infection. The present sequence was used to  
CC illustrate the invention.

XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;

SQ

Query Match 1.2%; Score 52.4; DB 8; Length 2000;  
Best Local Similarity 10.2%; Pred. No. 0.00026;  
Matches 103; Conservative 448; Mismatches 449; Indels 10; Gaps 4;

OY 2060 CTCACCATGATGACACCAAGACGATCTTCTTCACTTCAGGACATGCGGCAAGTGA 2119  
DB 12 CSSWGRRRMYMAGWSSCARWSSSRMSKMSGMSYKRCSSCGCKMTTRRSKYSSS 71  
OY 2120 TGAATGACGCTGTAGCCGCGATGTCGCAAGCGTTCGCAAGATGATGCTTCG 2179  
DB 72 ASGRTGSKWSSGSSYSKMKKRYKSKRMGRGRGMRSMRMRGRRCARSRMAG 131  
OY 2180 GTTGGGTGATGTGCTCTCTCTG-TGCAAGGCGCGTTGTTGCGGCAAGGAATGTG 2238  
DB 132 SGRMMGGSKRMSYMMWCYARGCGSCRKRSKXGSGKTCRGARAGSGSGAKYKSGSM 191  
OY 2239 GTTATCGACGTTGACGACAGCGGATGAGATTGATGACGACATACCCGTCGACGCGC 2298  
DB 192 SKRMMSSCGSGCGRRSAVSRYTSTKRYGTYKMTTYSASRCRAVTTTYSACGSY 251  
OY 2299 AAGTGGGAATGCTGCTGTGATTAACATCAACCTTCCGGAAGTCCGAGCGTGCT 2358  
DB 252 TWCSSKRSKMSMMKMKRMSRSYMSYSMMCTAYKKSYSRMCYMRGGGWRAT 311  
OY 2359 TCTGAATGAAAGTCAG--CTGCTCTCAACCTTACAGCTGTTACTCTGGAAGATGAG 2416  
DB 312 RYMGGRYMRMAMMYKMYRYRGYKMKRGMWAGMMRSMCRWAKACYMRMRMTR 371  
OY 2417 CCAGGGAAGAAGTGAAGTGGCCGCAATCGGTGATCGCTTATCAACATTTGCAAC 2476  
DB 372 RRRNAKSSRTSRKRRKMKRKYKMRGYSRRRSCKRARMMRCSGRMKWKGCGC 431  
OY 2477 GACAGTTAGCGAGCAAAAGCATGCTTGAATCGCCGATGATGTTGCGCAATGGT 2536  
DB 432 MTCRMSKSGMMRMSKWKMSMYKMSRMYRMRKCKCCTRTTMMGTBRGMMOTMKRCYK 491  
OY 2537 ACAGAGATTATTCAGTCGCGCACTCGAGTCAGCGGCGCGTGTGCTTTCTTCTG 2596  
DB 492 KRSKMKRKRMRMRMRMYRMRKRYMSARLYMRCAKKSYSASARARCYRRKGYW 551  
OY 2597 GGCTACCTGACAGTGAAGGAG--ACACTGAATCTTGTATGAAGTGGGTTGATC 2653  
DB 552 AGMMMKRKRMRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYRMYR 611  
OY 2654 CCAAGAACCAAAAGTACTTGCAAGCATCGCTGGATCTTCAAAAGCGGAATGTGACA 2713

DB 612 KSNRSRKRCKASKRSASAKRYAMMGATSGSRMSKSYTCWRKMSKSTCTMYYM 671  
OY 2714 CGTTGAAGTCCAACTTAACATCCGTGCG---GTGATAGATACATTACATGAT 2769  
DB 672 SKYTYAAYGSYMRRYRAMCWMMYMYRYSYRSTYMTAMTSTRTAMTGKISGRWT 731  
OY 2770 GCCGATTTGCGGGTGTGCTTGAAGAAATAGATGATCGGATTTCTCTCAAGTTC 2829  
DB 732 SWYTCCKSMKRYMRMYMSMMWAKTWKMRRYTRMMMYRYSKMYTCTMMGYWY 791  
OY 2830 AGACAGAGAGAGAGCTCTTTTACACTCTATCGAGCTGTGATGCTTCGCGCATCC 2889  
DB 792 WMTRTYMRMYMYKCTTYWWSATYWTGTAAAMWAKTRMKRMGTGAKTRGRARKAYW 851  
OY 2890 CCAGCCATTTCTCTAGTGAATCCAAAGGTTGAGACAGTCTCAAGCCAGACTCAC 2949  
DB 852 WKMTATCATKMMWKGAAGAKAMWAKAMKYYWSMRAYYYKTRRTYKTCMMVAR 911  
OY 2950 AGTCTCAAGATGTAATCTTCTTACTAATAAGAGATGACCGCTTGCTAAGAGCTA 3009  
DB 912 WGSMAVRRMMSKSAKMMWMMKSGRMGWTKYVYCTTMMACGRATKYMCCAGMAYVS 971  
OY 3010 TCTGTGAGACTYACGACGCGGATATGCGCTGCGTCTGCTGGATCCGGA 3059  
DB 972 YSMTRTYMRTWRMWMWASSRTAKRMAMMMKTRAMSKSYAAAYWKMAACA 1021

RESULT 6  
ADA71938/C  
ID ADA71938 standard; DNA; 2000 BP.

XX ADA71938;  
XX  
XX 20-NOV-2003 (first entry)  
XX  
XX Rice gene, SEQ ID 5263.  
XX  
XX Plant; bacterial infection; fungal infection; viral infection; rice;  
XX gene; ds.  
XX  
XX Oryza sativa.  
XX  
XX MO200300898-A1.  
XX  
XX 03-JAN-2003.  
XX  
XX 22-JUN-2001; 2001WO-IB001105.  
XX  
XX 22-JUN-2001; 2001WO-IB001105.  
XX  
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
XX Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
XX WPI; 2003-175290/17.

DR Identifying at least one gene involved in plant resistance or response to  
PT pathogenic infection for conferring resistance or tolerance to a plant to  
PT bacterial, fungal or viral infection by determining or detecting plant  
PT gene expression.

PS Claim 27; SEQ ID NO 5263; 899pp; English.

XX The present invention relates to a method (M1) for identifying genes  
CC involved in plant resistance or response to pathogenic infection. M1  
CC comprises identifying a gene whose expression is significantly altered in  
CC the incompatible interaction of plant gene expression relative to  
CC expression of the gene in an uninfected plant, in a mutant plant that  
CC does not express a gene associated with response to pathogenic infection,  
CC or in a corresponding incompatible or compatible interaction. (M1) is  
CC useful for conferring resistance to resistance or tolerance to a plant to  
CC bacterial, fungal or viral infection. The present sequence was used to

CC illustrate the invention.  
XX Sequence 2000 BP; 336 A; 265 C; 284 G; 363 T; 0 U; 752 Other;  
SQ

Query Match  
Best Local Similarity 1.2%; Score 51.8; DB 8; Length 2000;  
Matches 71; Conservative 309; Mismatches 301; Indels 4; Gaps 2;

QY 1454 TTGAGACCAAGTAGTGGCCGTTGTTCTCTGCTGTTTAGACTACATCCGACAACCT 1513  
DB 711 WTKRAKSYRBRMYRMWYKGMFTYRMYRSCMTBARMSRRKMGASMSKMMYRGA 652  
QY 1514 CGCCTACTGCGCCCTTACTCTTGAAGCTGAAGCCGCTCATTTGAGAGGCGCTGTC 1573  
DB 651 RSMWYSKSCSAKCCCKTTRYTSSYMSYKSMWSMTSKMSMGMTCTMTSM 592  
QY 1574 GACTCACCCTGCGTTCGCTGCTGATAGTGTTCGAGAT---CCTATACCGTCGCTTA 1630  
DB 591 KGTBRBSKMGMSMSMRMYRMWYKMKRKRKMYRMKCTBRMCYTRMGTYTTSRSR 532  
QY 1631 CGAGCACCAGCCCAAGTGTACCGCGGTGTGACAGCAACAACAGGTGCGGTGAGAGAAG 1690  
DB 531 MMYTGARYKARYTSRRYMMYKRYKCYWYYGYMYMKGSYMMRYGYCAKCKCCYAMCMA 472  
QY 1691 TCATCCAGTGGCTCAGCATGGGGCAACATTTCTTTGAGCGCCGCAATGGCGGCTTCT 1750  
DB 471 AYSMMMYWYRKYKMMWMSMTKYMSWYKCRSMKYGAKCYGCKMYTCSYGMKMYT 412  
QY 1751 TCGCCAAGATGCCGATACAGAAACCTCTCAGAGAGTTCCAGCTCCGCGCGAGGACC 1810  
DB 411 YMSYKYSRCKYKWM-RMYMYGMMYMYYSAYSSMTMYYYAKYMYKRYKRRTGMYG 353  
QY 1811 CGAAACCATCATCAAGAGAGAGTCACTTCTTTGCGGAGACCGGATTAAGTTCCGAC 1870  
DB 352 KSYKKKCTCWMYKCMKRCYRMKMRKTKYKSRKCYMRATCYMCCYRKGRMYRRS 293  
QY 1871 CTGATGCTTCAACAACAGATCTCTCGTCCGAGAGAAACCTGTAGAGCAAGACCG 1930  
DB 292 MMRTAGMKRMSWMSRWCYSWYMYKMKWKSYSYMSYGMARSSGYSRSAAKRTYKYS 233  
QY 1931 AGTTCAAGTTAGTCAATGCTGACTGCTGCTCAACTGCAACAACAACATTGGCAGC 1990  
DB 232 TSTRAKMRARCMYASACRYSRTSYCGCSGSSKKYKMSKSMMTGSSWSCCYTC 173  
QY 1991 CGCACTCAAGTTGTTCTCCCGTATCCAGCTCGGTGAGTAGACATATGCAATTATGA 2050  
DB 172 YGCAWMSGCMWYMSGCGYTRGMKMRKYSKCKYSCCTKYCSYGYRYCKMYK 113  
QY 2051 CATTGAGCGCTCAGACATGAGACACCAAGACCGATCTTCTTCACTTCAAGCACTG 2110  
DB 112 YSYKCYCYCYWYMSYMRMYMKCMCRSCSSWMSCAYSTSTSRWMSMYAAKMGCG 53  
QY 2111 GCGAAGTAGATGATGAGCGGTAGG 2135  
DB 52 SSGMYRMSKSKMYSKYSCKYTS 28

RESULT 7  
AA03159  
ID AA03159 standard; DNA; 1164 BP.  
XX  
AC AA03159;  
XX  
XX 25-MAR-2003 (revised)  
DT 02-FEB-1991 (first entry)  
XX  
DE Complete gene sequence of wildtype (WT) Streptomyces murinus glucose  
isomerase (GI).  
XX  
XX Site-directed mutagenesis; glucose isomerase; high fructose corn syrup;  
KM 89.  
XX  
OS Streptomyces murinus.

XX Key Location/Qualifiers  
FH CDS 1..1164  
FT CDS /\*tag= a  
XX  
XX  
XX EP351029-A.  
PN 17-JAN-1990.  
XX  
XX 17-JUL-1989; 89EP-00201892.  
XX  
XX 15-JUL-1988; 88EP-00201539.  
PR 04-NOV-1988; 88EP-00402789.  
PR 17-JUL-1989; 89EP-00201893.  
XX  
XX (KONN ) GIST-BROCADES NV.  
PA (PLBZ ) PLANT GENETIC SYSTEMS NV.  
PA (GEMV ) GENENCOR INT INC.  
XX  
XX Luiten RGM, Quax WJ, Schuurhuiz FW, Mrabet N;  
PI  
XX  
XX WPI; 1990-016368/03.  
DR P-PSDB; AAR02215.  
XX  
XX  
PT New mutant glucose isomerase enzymes - obtd. by site-directed mutagenesis  
of gene from Actinoplanes missouriensis, used for prodn. of high fructose  
corn syrups.  
PT  
XX  
XX Example; Fig 19; 56pp; English.  
XX  
XX The invention is a novel mutant glucose isomerase (GI) enzyme with  
improved properties thru AA substn. The GI is pref. derived from  
CC Actinoplanes missouriensis. The substn. is lys for Arg. or vice versa.  
CC Specific lys or Arg residues are identified and site-directed  
CC mutagenesis of the DNA sequence encoding the GI is performed. The novel  
CC mutant gi shows 65 per cent or more sequence homology with the AA  
CC sequence of WT A. missouriensis GI. Compared to wt enzyme it shows higher  
CC conversion performance and improved thermostability and pH stability. The  
CC example concerns the cloning and sequencing of GI from other bacterial  
CC strains. Fig 21 displays the AA S0s to demonstrate the mutual homology.  
CC (Updated on 25-MAR-2003 to correct PR field.) (Updated on 25-MAR-2003 to  
CC correct PA field.)  
XX  
SQ Sequence 1164 BP; 178 A; 445 C; 382 G; 159 T; 0 U; 0 Other;  
Query Match  
Best Local Similarity 1.2%; Score 51.6; DB 2; Length 1164;  
Matches 150; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 31 AGCCCGTCGAGGAAATCATAAACGGCTCAATACGACTACAACTGGGCTTCCAGTGT 90  
DB 16 ACCCCGAGGACAGGTTCACTTCGCTGTGTGACCGTGGCTGGAGGAAAGGACCCG 75  
QY 91 GTGCGAGACACACTCTCAACCCCGCAACCGCGGAAGAGTGGCCGAGTAGCGAGAT 150  
DB 76 TTGAGGAGCGCACCGCCCGCTCGAACCGGTGAGAGCGGTGACGCGCTGGCGAG 135  
QY 151 TTGCGTGCATGACAAGATCTACAGAGCCCTGAATTTCTTACTGCGGGAAGATGAC 210  
DB 136 CTGGGCGCTTACGGAATGACTTTCACAGAGAGAGACTATCCCTTCGGGTCTCCGAC 195  
QY 211 TCCCTTAACGAGCAGAGCCAACTTCTTCATGAGAGCCAAAGCTGCGAGCTGAAGTGG 270  
DB 196 ACCGACGCGAGTCGACATCAAGCCCTTCCGCAAGCCCTGGACCGCACCGGCAATGACG 255  
QY 271 GTGCGCAAGGCCAGCGCCGACCTGACAGGCTTCGCTGATCCAGAGAACCTCCCGCGCC 330  
DB 256 GTGCCATGCGCCACCAACCACTTACCCACACCCGCTTCAAGGAGCGGCGCTTACCC 315  
QY 331 GCTACTGCGGCGCA 344  
DB 316 GCCAAGACCGCGCA 329



```
RESULT 8
AAS17845
ID AAS17845 standard; cDNA, 3731 BP.
XX
XX AAS17845;
XX
XX 08-MAY-2002 (first entry)
XX
XX Tomato RNA-directed RNA polymerase (RdRP) cDNA from clone HF.
XX
XX Tomato; RdRP; RNA-directed RNA polymerase; in vitro transcription;
XX cytosolic; virucide; RNA synthesis inhibitor; antibody; immunogen;
XX transgenic plant; transgenic animal; cancer; viral infection;
XX immunoprecipitation; immunolocalisation; RdRP/HF; mutant; gene therapy;
XX ss.
XX
XX Lycopersicon esculentum.
XX
XX Key Location/Qualifiers
XX mutation replace(31,T)
XX mutation /*tag= b
XX mutation replace(32,G)
XX CDS /*tag= c
XX /*tag= 194..3535
XX /*product= "RdRP protein"
XX /*EC_number= "2.7.7.48"
XX mutation replace(2017,G)
XX mutation /*tag= d
XX mutation replace(2264,T)
XX mutation /*tag= e
XX mutation replace(3597,A)
XX mutation /*tag= f
XX mutation replace(3599,A)
XX /*tag= g
XX
XX US2001023067-A1.
XX
XX 20-SEP-2001.
XX
XX 08-FEB-2001; 2001US-00782874.
XX
XX 05-MAR-1997; 97US-00811583.
XX
XX (WASS/) WASSENEGGER M.
XX (RIED/) RIEDEL L.
XX
XX Wassenegger M, Riedel L, Schiebel W, Sanger HL;
XX
XX WPI: 2001-595798/67.
XX
XX New nucleic acid molecule encoding a polypeptide having the enzymatic
XX activity of RNA-directed RNA polymerase, for modulating gene expression
XX and treating cancer and virus infection in human and animals.
XX
XX Example 3; Page; 34pp; English.
XX
XX This sequence represents a cDNA encoding the tomato RNA-directed RNA
XX polymerase (RdRP/HF) of the invention. The invention comprises the
XX nucleic acid and protein sequences of RdRP. The protein of the invention
XX can catalyse in vitro transcription of short single stranded RNAs into
XX cDNA molecules, this transcription can be either primed by RNA or DNA
XX oligonucleotides or be unprimed. The protein may have cytosolic or
XX virucide activities. The sequences of the invention may be used in gene
XX therapy or as an RNA directed RNA synthesis inhibitor. The RdRP cDNA
XX sequence and a template nucleic acid molecule derived from a nucleic acid
XX molecule which causes a disease are useful for treating a disease caused
XX by the undesired expression or overexpression of a nucleic acid molecule
XX in a human, rat or mouse, by administering the molecules. This system can
XX be used in the preparation of a pharmaceutical composition and for
XX inhibiting expression of any desired gene by transferring the RdRP system
XX to organisms that either lack a comparable mechanism or do not
```

```
CC sufficiently express their own RdRP. An antibody or an antagonist or
CC inhibitor to the protein are useful for inhibiting RNA directed RNA
CC synthesis and for ensuring stable heterologous, gene expression in
CC transgenic organisms. The sequence is useful for probes and/or for the
CC control of gene expression, as primers for amplification of nucleic acid
CC molecules and as tools for the detection of expression of the cDNA
CC molecules. Additionally, nucleotide and protein sequences are useful for
CC suppression of undesired gene expression in humans and animals. The RdRP
CC is useful as a therapeutic agent for the control of cancer and virus
CC infection in humans and animals and the antibody is useful for
CC immunoprecipitation or immunolocalisation of the protein, identification
CC of polypeptides interacting with it and screening expression libraries.
CC Note: This sequence is not shown in the specification but was created by
CC the indexer from the wild type RdRP sequence shown in AAS17837 and the
CC information shown in table 1
XX
XX SQ Sequence 3731 BP; 1062 A; 672 C; 848 G; 1149 T; 0 U; 0 Other;
XX
XX Query Match 1.2%; Score 49; DB 4; Length 3731;
XX Best Local Similarity 51.6%; Pred. No. 0.0043;
XX Matches 112; Conservative 0; Mismatches 105; Indels 0; Gaps 0;
XX
XX QY 2850 TACACTCCATGCGACTGTGATGTCCTCGTGCCGATGCCAGCCCAATTCCTAGTGA 2909
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX Db 2425 TAATTTCATTCTGAAGGGAATGTGTTGCAAAAATTCATGTCATCCTGGTGA 2484
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX QY 2910 TATCCACGGGTTGACGACGCTTCAAGCCAGACCTCCACAGTCTCAAGATGATCAT 2969
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX Db 2485 TATTCGTGTTTAAAGCCTGTAATGTTTCAGCCCTCACCACATGGTAGATTGTGTGT 2544
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX QY 2970 CTTCTCTACTAAGAGAGATGACCGCTTGCTAAGACCTATCTGTGAGACTACGACGG 3029
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX Db 2545 ATTCCCTCAGAAAGAAAGAAAGACCTCATCCGAATGATTTCTGGAGTATTTGATGG 2604
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX QY 3030 CGATATGAGCTTGAGTCTGCTGGGATCCGAGATCTGC 3066
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX Db 2605 GGATATCTACTTGTGTTGCTGGGATCAAGACATGATC 2641
XX ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
XX
XX RESULT 9
XX AAS17837
XX ID AAS17837 standard; cDNA; 3731 BP.
XX
XX AC AAS17837;
XX
XX DT 08-MAY-2002 (first entry)
XX
XX DE Tomato RNA-directed RNA polymerase (RdRP) cDNA.
XX
XX KW Tomato; RdRP; RNA-directed RNA polymerase; in vitro transcription;
XX cytosolic; virucide; RNA synthesis inhibitor; antibody; immunogen;
XX transgenic plant; transgenic animal; cancer; viral infection;
XX immunoprecipitation; immunolocalisation; ss; gene therapy.
XX
XX Lycopersicon esculentum.
XX
XX Key Location/Qualifiers
XX CDS /*tag= a
XX /*tag= 194..3538
XX /*product= "RdRP protein"
XX /*EC_number= "2.7.7.48"
XX
XX US2001023067-A1.
XX
XX 20-SEP-2001.
XX
XX 08-FEB-2001; 2001US-00782874.
XX
XX 05-MAR-1997; 97US-00811583.
XX
XX (WASS/) WASSENEGGER M.
XX (RIED/) RIEDEL L.
XX
```

PI Massenegger M, Riedel L, Schiebel W, Sanger HL;  
 XX WPI; 2001-595798/67.  
 DR P-PSDB; AAU10006.  
 XX  
 PT New nucleic acid molecule encoding a polypeptide having the enzymatic  
 PT activity of RNA-directed RNA polymerase, for modulating gene expression  
 PT and treating cancer and virus infection in human and animals.  
 PS Claim 1; Page 15-19; 34pp; English.  
 XX  
 XX This sequence represents a cDNA encoding the tomato RNA-directed RNA  
 CC polymerase (RdRp) protein of the invention. The invention comprises the  
 CC nucleic acid and protein sequences of RdRp. The protein of the invention  
 CC can catalyze in vitro transcription of short single stranded RNAs into  
 CC DNA molecules, this transcription can be either primed by RNA or DNA  
 CC oligonucleotides or be unprimed. The protein may have cytosolic or  
 CC virucide activities. The sequences of the invention may be used in gene  
 CC therapy or as an RNA directed RNA synthesis inhibitor. The RdRp cDNA  
 CC sequence and a template nucleic acid molecule derived from a nucleic acid  
 CC molecule which causes a disease are useful for treating a disease caused  
 CC by the undesired expression or overexpression of a nucleic acid molecule  
 CC in a human, rat or mouse, by administering the molecules. This system can  
 CC be used in the preparation of a pharmaceutical composition and for  
 CC inhibiting expression of any desired gene by transferring the RdRp system  
 CC to organisms that either lack a comparable mechanism or do not  
 CC sufficiently express their own RdRp. An antibody or an antagonist or  
 CC inhibitor to the protein are useful for inhibiting RNA directed RNA  
 CC synthesis and for ensuring stable heterologous gene expression in  
 CC transgenic organisms. The sequence is useful for probes and/or for the  
 CC control of gene expression, as primers for amplification of nucleic acid  
 CC molecules and as tools for the detection of expression of the cDNA  
 CC molecules. Additionally, nucleotide and protein sequences are useful for  
 CC suppression of undesired gene expression in humans and animals. The RdRp  
 CC is useful as a therapeutic agent for the control of cancer and virus  
 CC infection in humans and animals and the antibody is useful for  
 CC immunoprecipitation or immunolocalisation of the protein, identification  
 CC of polypeptides interacting with it and screening expression libraries  
 XX  
 SQ Sequence 3731 BP; 1064 A; 669 C; 849 G; 1149 T; 0 U; 0 Other;  
 Query Match 1.2%; Score 49; DB 4; Length 3731;  
 Best Local Similarity 51.6%; Pred. No. 0.0043;  
 Matches 112; Conservative 0; Mismatches 105; Indels 0; Gaps 0;  
 QY 2850 TACACTCTTATCGAGCTGATGTCCTCGTGGCGCGATCCCGACCCATTTCCTAGTGA 2309  
 Db 2425 TAATTTTCATTCGAAAGGAAATGTTGTCAAAAAATTCATGCTTGATCTGGTGA 2484  
 QY 2910 TATTCACAGGGTTCGACACTCTTCAAGCCAGAGCTCCAGATGTCAGAGATGAATCAT 2369  
 Db 2485 TATTCGTTGTTTAAAGGCTGTAATGTTGAGCGCTGCACACATGTAATGTTGT 2544  
 QY 2970 CTTCCTACTAAAGAGATGATACCGCTTGCTAAGAGCTATCTGGTGAAGACTACGACGG 3029  
 Db 2545 ATTCCCTCAGAAAAGAAAAGACCTCATCCGAATGAATGTTCTGGAGATGATTTGGATGG 2604  
 QY 3030 CGATATGCGCTGGGCTGCTGGGATCCGAGATCGTC 3066  
 Db 2605 GGATATCTACTTTGTTGCTGGGATCAAGACATGATC 2641  
 RESULT 10  
 AAD04370  
 ID AAD04370 standard; cDNA; 3731 BP.  
 XX  
 XX AAD04370;  
 XX  
 DT 04-JUL-2001 (first entry)  
 XX  
 XX Tomato RNA-directed RNA polymerase (RdRp) cDNA.  
 DE Tomato; gene therapy; RNA-directed RNA polymerase; RdRp; gene expression;  
 XX  
 KM

KW transgenic plant; tissue culture; plant breeding; therapy; C-protein; ss.  
 XX  
 XX Lycopersicon esculentum.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 194..3538  
 FT /tag= a  
 FT /product= "Tomato C-protein having RNA-directed RNA  
 FT polymerase (RdRp) activity"  
 XX  
 XX US6218142-B1.  
 XX  
 PD 17-APR-2001.  
 XX  
 PF 05-MAR-1997; 97US-00811583.  
 XX  
 PR 05-MAR-1997; 97US-00811583.  
 XX  
 PA (WASS/) MASSENEGGER M.  
 PA (RIED/) RIEDEL L.  
 PI Massenegger M, Riedel L, Schiebel W, Sanger HL;  
 XX WPI; 2001-289830/30.  
 DR P-PSDB; AAE00897.  
 XX  
 XX New nucleic acid molecules encoding polypeptides with RNA-directed RNA  
 PT polymerase enzymatic activity, useful in modulating gene expression in  
 PT plants, humans and animals, as well as in plant cell/tissue cultures or  
 PT plant breeding.  
 XX  
 PS Claim 1; Col 25-34; 31pp; English.  
 XX  
 XX The present sequence is a cDNA encoding tomato C-protein having RNA-  
 CC directed RNA polymerase (RdRp) activity. This protein is capable of RNA-  
 CC directed RNA synthesis, thus using RNA as a template for synthesising  
 CC complementary RNA molecules. RdRp nucleic acid is useful for modulating  
 CC gene expression in plants, humans and animals. This may lead to various  
 CC physiological, developmental and/or morphological changes. Transgenic  
 CC plants containing RdRp nucleic acid is especially useful in plant cell or  
 CC tissue cultures and in plant breeding. RdRp is useful in gene therapy,  
 CC particularly for treating a disease that is caused by the undesirable  
 CC expression or overexpression of a gene  
 XX  
 SQ Sequence 3731 BP; 1064 A; 669 C; 849 G; 1149 T; 0 U; 0 Other;  
 Query Match 1.2%; Score 49; DB 4; Length 3731;  
 Best Local Similarity 51.6%; Pred. No. 0.0043;  
 Matches 112; Conservative 0; Mismatches 105; Indels 0; Gaps 0;  
 QY 2850 TACACTCTTATCGAGCTGATGTCCTCGTGGCGCGATCCCGACCCATTTCCTAGTGA 2309  
 Db 2425 TAATTTTCATTCGAAAGGAAATGTTGTCAAAAAATTCATGCTTGATCTGGTGA 2484  
 QY 2910 TATTCACAGGGTTCGACACTCTTCAAGCCAGAGCTCCAGATGTCAGAGATGAATCAT 2369  
 Db 2485 TATTCGTTGTTTAAAGGCTGTAATGTTGAGCGCTGCACACATGTAATGTTGT 2544  
 QY 2970 CTTCCTACTAAAGAGATGATACCGCTTGCTAAGAGCTATCTGGTGAAGACTACGACGG 3029  
 Db 2545 ATTCCCTCAGAAAAGAAAAGACCTCATCCGAATGAATGTTCTGGAGATGATTTGGATGG 2604  
 QY 3030 CGATATGCGCTGGGCTGCTGGGATCCGAGATCGTC 3066  
 Db 2605 GGATATCTACTTTGTTGCTGGGATCAAGACATGATC 2641  
 RESULT 11  
 ACH16195  
 ID ACH16195 standard; cDNA; 495 BP.  
 XX  
 XX ACH16195;  
 AC  
 AC ACH16195;  
 XX

DT 13-OCT-2003 (first entry)  
 XX  
 DE Human adult heart cDNA #509.  
 XX  
 KW Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;  
 KW genome mapping; biodiversity; genetic disorder.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003073623-A1.  
 XX  
 PD 17-APR-2003.  
 XX  
 PF 30-JUL-2001; 2001US-00918995.  
 XX  
 PR 30-JUL-2001; 2001US-00918995.  
 XX  
 PA (DRMA/) DRMANAC R T.  
 PA (LABA/) LABAT I.  
 PA (STAC/) STACHE-CRAIN B.  
 PA (DICK/) DICKSON M C.  
 PA (JONE/) JONES L W.  
 PI Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;  
 PI WPI; 2003-615964/58.  
 DR  
 XX  
 PT New polynucleotide sequences obtained from various cDNA libraries, useful  
 PT as hybridization probes, as oligomers for PCR, for chromosome and gene  
 PT mapping, in the recombinant production of protein, or in generating  
 PT antisense DNA or RNA.  
 XX  
 PS Claim 1; SEQ ID NO 3407; 44pp; English.  
 XX  
 CC The invention relates to an isolated polynucleotide comprising any one of  
 CC 38043 CDNA sequences, appearing as ACH12789-ACH50831, whose sequence was  
 CC determined by the technique of SBH (sequencing by hybridisation). Also  
 CC included is a purified polypeptide comprising a sequence corresponding to  
 CC a reading frame of the novel polynucleotide. The nucleic acid sequences  
 CC are useful in diagnostics as expressed sequence tags (EST) for  
 CC identifying expressed genes or for physical mapping of the human genome,  
 CC in forensics, in assessing biodiversity, or in identifying mutations  
 CC responsible for genetic disorders and other traits. The nucleotide  
 CC sequences are also useful as hybridisation probes, as oligomers for PCR,  
 CC for chromosome and gene mapping, in the recombinant production of  
 CC protein, or in generating antisense DNA or RNA. The purified polypeptide  
 CC is useful for generating antibodies specific for it. The present sequence  
 CC is one of the 38043 isolated CDNA/EST sequences. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?DocID=20030073623  
 CC  
 XX  
 SQ Sequence 495 BP; 137 A; 123 C; 123 G; 111 T; 0 U; 1 Other;  
 Query Match 1.1%; Score 45.6; DB 9; Length 495;  
 Best Local Similarity 51.5%; Pred. No. 0.013; 99; Indels 0; Gaps 0;  
 Matches 105; Conservative 0; Mismatches 99; Indels 0; Gaps 0;  
 QY 2506 AATGCCCAAGTGAATTTGGCAATGGTTTACGAGATTTTCAGTCGCGCAACTGCA 2565  
 DB AATACAGAGTGGCTTTCTTCTACTGCAGATGATGACATATCAATGCCCAACAAA 346  
 QY 2566 GTACAGCCAGCGCGGTGCTTTTCTTCTGGGCTACTGACAGTCAGAGAGACACTG 2625  
 DB GACCTCTTGACCCCTTGATCTTGGCTCTGGAAAGAGACAGCAAGCAATACCTGAA 406  
 QY 2626 AACTCTTGATGAACAGTGGCTTCATCCCAAGAAAGCAAAAGTACTTCCAGACATCGCC 2685  
 DB CTCCTCCGATGAACCCGTACGTCAACAGGCGCTGAAAAACAACTTGGAAAGAACTGCA 466  
 QY 2686 TGGGATCTTCAAAAGCGGAAATGT 2709  
 DB 467 TTGTATTTGCCAGAGGACAAGT 490

RÉSULT 12  
 AAC84447  
 ID AAC84447 standard; CDNA; 2710 BP.  
 AAC84447;  
 AC  
 XX  
 DT 02-APR-2001 (first entry)  
 XX  
 DE Nucleotide sequence of mdt cDNA clone ID No: 237703.2.  
 XX  
 KW Molecule for disease detection and treatment; mdt; human; cirrhosis;  
 KW cell proliferative disorder; arteriosclerosis; bursitis; hepatitis;  
 KW connective tissue disease; myelofibrosis; cancer; leukemia; autoimmune;  
 KW acquired immunodeficiency syndrome; AIDS; Addison's disease; allergy;  
 KW anemia; bronchitis; gout; Hashimoto's thyroiditis; multiple sclerosis;  
 KW cytostatic; immunomodulator; anti-inflammatory; gene therapy; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200075298-A2.  
 XX  
 PD 14-DEC-2000.  
 XX  
 PF 01-JUN-2000; 2000WO-US015344.  
 XX  
 PR 03-JUN-1999; 99US-0137412P.  
 PR 05-AUG-1999; 99US-0147500P.  
 PR 05-AUG-1999; 99US-0147501P.  
 PR 05-AUG-1999; 99US-0147542P.  
 XX  
 PA (INCY-) INCYTE GENOMICS INC.  
 XX  
 PI Hodgson DM, Lincoln SE, Russo FD, Spiro PA, Banville SC,  
 PI Bratcher SR, Dufour GF, Cohen HJ, Rosen BH, Chalup MS, Hillman JL,  
 PI Jones AL, Yu YV, Greenwalt LB, Panzer SR, Roseberry AM, Wright RJ,  
 PI Daniels SE;  
 PI WPI; 2001-071068/08.  
 DR  
 XX  
 PT New polypeptide and polynucleotide molecules for disease detection and  
 PT treatment are useful in diagnosis and gene therapy of proliferative  
 PT disorders (e.g. breast cancer), autoimmune and inflammatory disorders  
 PT (e.g. AIDS or allergy).  
 XX  
 PS Claim 1; Page 90; 99pp; English.  
 XX  
 CC Sequences AAC84445-C84458 are new isolated polynucleotide molecules for  
 CC disease detection and treatment (mdt). The mdt polynucleotides are  
 CC useful for diagnosing or treating disorders associated with disease  
 CC detection and treatment molecules. These diseases include cell  
 CC proliferative disorders (e.g. arteriosclerosis, bursitis, cirrhosis,  
 CC hepatitis, mixed connective tissue disease, myelofibrosis, or cancers  
 CC such as leukemia, or breast or brain cancers), autoimmune or inflammatory  
 CC disorders (e.g. acquired immunodeficiency syndrome (AIDS), Addison's  
 CC disease, allergies, anemia, autoimmune haemolytic anemia, bronchitis,  
 CC gout, Hashimoto's thyroiditis, multiple sclerosis or cancers). The mdt  
 CC are especially useful for somatic or germline gene therapy. The mdt may  
 CC also be used to detect the presence of, or to quantify the amount of, an  
 CC mdt-related polynucleotide in a sample. The mdt are also useful for  
 CC isolating full length cDNA sequences utilizing hybridization and/or  
 CC amplification procedures, for generating hybridization probes useful in  
 CC chromosomal mapping of naturally occurring genomic sequences, as  
 CC molecular weight markers, or for monitoring the progress of disorders  
 CC associated with abnormal levels of mdt expression or evaluating the  
 CC efficacy of a particular treatment. The recombinant nucleic acids are  
 CC useful as part of a viral vector (e.g. based on a vaccinia virus) for  
 CC vaccinating a mammal and inducing a protective immunological response in  
 CC the mammal. The MDT polypeptides are useful for screening molecules that  
 CC bind to or are bound by the polypeptides, i.e. for screening agonists,  
 CC antagonists or modulators of MDT

Sequence 2710 BP; 856 A; 516 C; 450 G; 880 T; 0 U; 8 Other;  
Query Match 1.1%; Score 45.6; DB 4; Length 2710;  
Best Local Similarity 51.5%; Pred. No. 0.039;  
Matches 105; Conservative 0; Mismatches 99; Indels 0; Gaps 0;  
QY 2506 AATGGCCAGTGGAAATTTCCGCAATGGGTTACAGAGTTATTCAGTCCGCACTCGA 2565  
DB 1010 AATCCAGAGTGGGTTCTTTCTTACTCAGCATATGCAATATCAATGCCCCAACAATA 1069  
QY 2566 GTCAGCCAGCGGCGTGTGCTTTCTTCTGGGCTGACATGCAAGAGAGACACTG 2625  
DB 1070 GGCCTCTTGAGCCCTCTTGATCTTGCTGTGGAAACAGACAGCAAGATACCTTGAA 1129  
QY 2626 AACTTCTTGATGAACACTGGGTTGATCCCAAGAGAAAAGTATCTGCAAGACATGCC 2685  
DB 1130 CTCCTCTGATGAACCGTTAGCTCAACAGGCGCTGAAAAACAATTGGAAGAACTGCA 1189  
QY 2686 TGGGATCTTCAAAAGCGGAATGT 2709  
DB 1190 TTGTATATTGCCAGAGACAAAGT 1213  
RESULT 13  
ID ABL71773 standard; cDNA; 282 BP.  
XX ABL71773;  
AC ABL71773;  
XX 14-MAY-2002 (first entry)  
DT XX  
DE Corn tassell-derived polynucleotide (cdps) SEQ ID NO:1147.  
XX  
XX Corn; corn tassell-derived polynucleotide; cdps; hybrid breeding; CDPS;  
KW inheritance; characteristic; growth; development; disease resistance;  
KW environmental adaptability; quality; yield; molecular marker;  
KW multigene trait; plant breeding; corn tassell; gene; ss.  
XX  
OS Zea mays.  
XX  
XX US2001051335-A1.  
PN 13-DEC-2001.  
PD  
XX 16-APR-1999; 99US-00294093.  
XX 21-APR-1998; 98US-0082567P.  
XX  
PA (LALG/) LALGUDI R V.  
PA (ITOL/) ITO L Y.  
PA (SHER/) SHERMAN B K.  
XX  
PI Lalgudi RV, Ito LY, Sherman BK;  
XX  
DR WPI; 2002-163647/21.  
XX  
PT Novel purified corn tassell-derived polynucleotide useful for determining  
PT altered gene expression, to recover regulatory elements and to follow  
PT inheritance of desirable characteristics through hybrid breeding  
PT programs.  
XX  
XX Claim 1; SEQ ID NO 1147; 201pp; English.  
XX  
XX The present sequence describes a purified corn tassell-derived  
XX polynucleotide sequence (cdps) comprising a nucleic acid sequence  
XX selected from those given in ABL70627 to ABL76833. The cdps sequences  
XX encode corn tassell-derived polypeptides (CDPS). The cdps sequences (1)  
XX can be used for determining altered gene expression, to recover  
XX regulatory elements and to follow inheritance of desirable  
XX characteristics through hybrid breeding programs. (1) are also useful in  
XX the evaluation, and alteration of desired characteristics associated with  
XX growth and development, disease resistance, environmental adaptability,  
XX quality and yield, and as molecular markers for studying inheritance of

CC multigene traits in a plant breeding program. (1) can be used to produce  
CC a tassell-specific profile of gene transcription, a transcript image, to  
CC clone regulatory elements for use in transformation vectors, to express a  
CC polypeptide, to identify, isolate or extend identical or related corn  
CC tassell nucleic acid sequences from DNA libraries, in nucleic acid  
CC hybridisation or amplification technologies, as query sequences to  
CC determine homology of known sequences, as probe for use in Southern or  
CC Northern hybridisation, and to identify the presence of and/or to  
CC determine the degree of similarity between two (or more) nucleic acid  
CC sequences  
XX  
SQ Sequence 282 BP; 78 A; 67 C; 70 G; 67 T; 0 U; 0 Other;  
Query Match 1.0%; Score 44; DB 6; Length 282;  
Best Local Similarity 52.1%; Pred. No. 0.029;  
Matches 98; Conservative 0; Mismatches 90; Indels 0; Gaps 0;  
QY 2867 GTGATGTCCTCGTGGGCGCATCCCGAGCCATTTCCCTAGTGTATTCACAGGGGTTGAG 2926  
DB 16 GTACTATCGTAATGCGAAGAAATCCATGCTTCATCAGGGAGTGTCCGATCTTGAAAG 75  
QY 2927 CAGTCTTCAAGCCAGAGCTCCACAGTCTCAAGGATGMAATCATCTTCTACTAAGAGAG 2986  
DB 76 CTGTGATGTCGCTGAACTGATCATCACCCTTGTATGCTTGTGCTTCCCAAGAAAGTG 135  
QY 2987 ATGTACCGCTTGTGAAGAACTATCTGTGTGAGACTACAGCGGCGATATGCGGTCT 3046  
DB 136 AGAGCCGCGACGCGAATGAAGCATCTGGAGTGATCTTGATGGGATATATCTTGTA 195  
QY 3047 GCTGGGAT 3054  
DB 196 CATGGGAT 203  
RESULT 14  
ID AAS33132 standard; cDNA; 1050 BP.  
XX AAS33132  
XX  
AC AAS33132;  
XX  
XX 04-DEC-2001 (first entry)  
DT XX  
DE DNA encoding human secreted protein, Seq ID No 91.  
XX  
XX Immunomodulatory; human immunodeficiency virus; HIV; anaemia; angina;  
KW rheumatoid arthritis; antiarteriosclerotic; cardiant; vascular;  
KW cerebroprotective; thrombolytic; antimicrobial; ophthalmological;  
KW cytostatic; Alzheimer's disease; Parkinson's disease; human; cancer;  
KW multiple sclerosis; cancer; hyperproliferative disorder; infection;  
KW Gaucher's disease; neurological disease; cerebrovascular disorder;  
KW thrombosis; wound healing; ss.  
XX  
OS Homo sapiens.  
XX  
XX  
XX WO200155326-A2.  
PN 02-AUG-2001.  
PD  
XX 17-JAN-2001; 2001WO-US001347.  
XX  
XX 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-018464P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216447P.  
PR 07-JUL-2000; 2000US-0216880P.



CC the production of antibodies and in assays to identify modulators  
 CC (agonists and antagonists) of the expression and activity of the secreted  
 CC proteins. The anti-(II) antibodies and antagonists may also be used to  
 CC down regulate expression and activity of (II). The anti-(II) antibodies  
 CC may also be used as diagnostic agents for detecting the presence of (II)  
 CC in samples (e.g. by enzyme linked immunosorbent assay (ELISA)). The  
 CC disorders include for example: immune/autoimmune diseases (e.g. HIV  
 CC (human immunodeficiency virus) infections, anaemia, rheumatoid arthritis  
 CC and multiple sclerosis), cancers and hyperproliferative disorders (e.g.  
 CC melanomas, neoplasms of the breast or liver, Sezary syndrome and  
 CC Gaucher's disease), neurological diseases (e.g. Alzheimer's disease,  
 CC Parkinson's disease and Charcot-Marie-Tooth disease), cardio-/  
 CC cerebrovascular disorders (e.g. cardiac arrest, tachycardia, angina and  
 CC thrombosis), infections caused by bacteria, viruses and fungi and ocular  
 CC disorders (e.g. corneal infections). (I) and (II), agonists, antagonists  
 CC and antibodies can also be used to promote wound healing, maintain organs  
 CC before transplantation, and support cell culture of primary tissues.

Query Match 1.0%; Score 44; DB 4; Length 1050;  
 Best Local Similarity 51.0%; Pred. No. 0.067;  
 Matches 104; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 2506 AATGCCAGTGGAAATTTCCCATATGGCTTACGAGATTATTCACATCCGCGCACTCGA 2565  
 DB 570 AATACCGAGAGTGGCTTTCTTTCTTACTGACATGATGCAATATCAATGCCAATAA 629  
 QY 2566 GTGAGCCAGCGCGGTGCTTTCTTCTGCTGGCTACCTGACATCAAGAGAGACACTG 2625  
 DB 630 GGCCTCTGAGCCCTTTCATCTTCTGCTGGAGACAGACAGCAAGATACCTTGAA 689  
 QY 2626 AACTCTTGATGAACAGTGGGTTGATCCCAAGAAAGAAAGTATGCAAGACATGCC 2685  
 DB 690 CTCCTCTGATGAACCGTTACGTCAACCAAGGCTGAAAAACAATTGGAAGAACTGCA 749  
 QY 2686 TGGGATCTTCAAAAGCGGAATGT 2709  
 DB 750 TTTGATATTGCCAGAGACAAGT 773

RESULT 15  
 AAF62769  
 ID AAF62769 standard; cDNA: 1318 BP.  
 XX AAF62769;

AC 03-MAY-2001 (first entry)  
 DT XX  
 DE Human FGIF cDNA.

XX Human; FGIF; foetal globin inducing factor; antihaemic; antianginal;  
 KW antiskling; analgesic; sickle cell anaemia; thalassaemia;  
 KM haemoglobinopathy; ss.

XX Homo sapiens.

XX US6184343-B1.

PN 06-FEB-2001.

PD 05-DEC-1997; 97US-00986304.

PF 06-DEC-1996; 96US-0033247P.

PR (STAM/) STAMATOYANNOPOLIOS G.  
 PA (PAPA/) PAPAYANNOPOULOU T.

XX Stamatoyanopoulos G, Papayannopoulou T;

XX WPI: 2001-225816/23.  
 DR P-P-SDB; AAB71861.

XX Novel fetal globin inducing factor capable of regulating expression of  
 PT globin gene, for treating globin disorders such as sickle cell anemia,

PT thalassemia and for screening modulators of fetal globin gene expression.  
 XX  
 PS Example; Fig 1; 16pp; English.

XX The present sequence encodes human foetal globin inducing factor (FGIF).  
 CC FGIF is useful for regulating the expression of globin genes such as  
 CC foetal globin. It may be used in the prophylactic or therapeutic  
 CC treatment of globin-associated disorders such as sickle cell anaemia and  
 CC thalassaemia. Symptoms associated with globin disorders treatable by FGIF  
 CC include anaemia, tissue hypoxia, organ dysfunction, abnormal haematocrit  
 CC values, ineffective erythropoiesis, abnormal reticulocyte count, abnormal  
 CC iron load, splenomegaly, hepatomegaly, impaired peripheral blood flow,  
 CC dyspnea, increased haemolysis, jaundice, anaemic crises and pain such as  
 CC angina pectoris. FGIF is also useful for screening for agents that  
 CC modulate its function. Agents that enhance FGIF function can be used to  
 CC treat haemoglobinopathies and beta-thalassaemia syndromes. A relatively  
 CC small number of FGIF molecules transactivates expression of endogenous  
 CC foetal globin genes. This method is highly advantageous over prior art  
 CC methods, using vectors carrying either human foetal or human adult globin  
 CC genes for expression of therapeutic polypeptides, which require a  
 CC consistent production of high levels of globin mRNA in the target cells  
 XX

SQ Sequence 1318 BP; 414 A; 245 C; 264 G; 395 T; 0 U; 0 Other;

Query Match 1.0%; Score 44; DB 5; Length 1318;  
 Best Local Similarity 51.0%; Pred. No. 0.077;  
 Matches 104; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 2506 AATGCCAGTGGAAATTTCCCATATGGCTTACGAGATTATTCACATCCGCGCACTCGA 2565  
 DB 592 AATACCGAGAGTGGCTTTCTTTCTTACTGACATGATGCAATATCAATGCCAATAA 651  
 QY 2566 GTGAGCCAGCGCGGTGCTTTCTTCTGCTGGCTACCTGACATCAAGAGAGACACTG 2625  
 DB 652 GGCCTCTGAGCCCTTTCATCTTCTGCTGGAGACAGACAGCAAGATACCTTGAA 711  
 QY 2626 AACTCTTGATGAACAGTGGGTTGATCCCAAGAAAGAAAGTATGCAAGACATGCC 2685  
 DB 712 CTCCTCTGATGAACCGTTACGTCAACCAAGGCTGAAAAACAATTGGAAGAACTGCA 771  
 QY 2686 TGGGATCTTCAAAAGCGGAATGT 2709  
 DB 772 TTTGATATTGCCAGAGACAAGT 795

Search completed: August 10, 2005, 19:57:15  
 Job time : 2065 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 10, 2005, 07:02:32 ; Search time 622 Seconds  
(without alignments)  
11064.598 Million cell updates/sec

Title: US-09-913-878A-1\_COPY\_2447\_6652

Perfect score: 4206

Sequence: 1 atgaccactactactcctagc.....tcacagggaatgcgcatrat 4206

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:\*

1: /cgn2\_6/ptodata/1/1na/5A\_COMB.seq:\*  
2: /cgn2\_6/ptodata/1/1na/5B\_COMB.seq:\*  
3: /cgn2\_6/ptodata/1/1na/6A\_COMB.seq:\*  
4: /cgn2\_6/ptodata/1/1na/6B\_COMB.seq:\*  
5: /cgn2\_6/ptodata/1/1na/PCITUS\_COMB.seq:\*  
6: /cgn2\_6/ptodata/1/1na/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	51.6	1.2	1164	1	Sequence 6, Appli
2	49	1.2	3731	3	Sequence 1, Appli
3	44	1.0	1318	3	Sequence 1, Appli
4	41	1.0	1929	4	Sequence 9129, Ap
5	41	1.0	10216	4	Sequence 976, App
6	40.8	1.0	4403765	3	Sequence 2, Appli
7	40.8	1.0	4411529	3	Sequence 1, Appli
8	40.2	1.0	803	4	Sequence 12025, A
9	39.6	0.9	1209	4	Sequence 1706, Ap
10	39.6	0.9	1509	4	Sequence 12165, A
11	39.6	0.9	1692	4	Sequence 12425, A
12	38.8	0.9	7218	1	Sequence 14, Appli
13	38.6	0.9	3192	4	Sequence 2923, Ap
14	38.6	0.9	3489	4	Sequence 2922, Ap
15	38.6	0.9	36093	4	Sequence 14664, A
16	38.6	0.9	36093	4	Sequence 14665, A
17	38	0.9	17897	4	Sequence 1182, Ap
18	37.2	0.9	505	4	Sequence 15639, A
19	36.8	0.9	1548	2	Sequence 5, Appli
20	36.8	0.9	1548	2	Sequence 5, Appli
21	36.8	0.9	1581	2	Sequence 6, Appli
22	36.8	0.9	1581	2	Sequence 6, Appli
23	36.8	0.9	2142	4	Sequence 13530, A
24	36.6	0.9	1065	4	Sequence 31, Appli
25	36.4	0.9	1160	4	Sequence 909, App
26	36.2	0.9	816	4	Sequence 11434, A
27	36.2	0.9	1332	4	Sequence 4685, Ap

28	36.2	0.9	1473	4	US-09-252-991A-4812	Sequence 4812, Ap
29	36.2	0.9	1524	4	US-09-252-991A-11455	Sequence 11455, A
30	36.2	0.9	1551	4	US-09-252-991A-4758	Sequence 4758, Ap
31	36.2	0.9	2028	4	US-09-252-991A-11416	Sequence 11416, A
32	35.4	0.8	1437	4	US-09-902-540-8578	Sequence 8578, Ap
33	35.4	0.8	1722	3	US-09-385-028-15	Sequence 15, Appli
34	35.4	0.8	1722	4	US-09-726-614-15	Sequence 15, Appli
35	35.4	0.8	1722	4	US-09-385-040-15	Sequence 905, App
36	35.4	0.8	9081	4	US-09-902-540-905	Sequence 13, Appli
37	35.4	0.8	11604	3	US-09-385-028-13	Sequence 13, Appli
38	35.4	0.8	11604	4	US-09-726-614-13	Sequence 13, Appli
39	35.4	0.8	11604	4	US-09-385-040-13	Sequence 1, Appli
40	35.4	0.8	15079	3	US-09-385-028-1	Sequence 1, Appli
41	35.4	0.8	15079	4	US-09-726-614-1	Sequence 1, Appli
42	35.4	0.8	15120	4	US-09-385-040-1	Sequence 1, Appli
43	35.2	0.8	1546	3	US-09-383-318A-1	Sequence 1, Appli
44	35.2	0.8	3711	4	US-09-902-540-7049	Sequence 7049, Ap
45	35.2	0.8	13624	4	US-09-902-540-1053	Sequence 1053, Ap

#### ALIGNMENTS

RESULT 1  
US-07-640-476-6  
; Sequence 6, Application US/07640476  
; Patent No. 5376536  
; GENERAL INFORMATION:  
; APPLICANT: QUAX, WILHELMUS  
; APPLICANT: LUTTEN, RUDOLF G.M.  
; APPLICANT: SCHURHUIZEN, PAUL W.  
; APPLICANT: MRABET, NADIR  
; TITLE OF INVENTION: NOVEL GLUCOSE ISOMERASE ENZYMES AND  
; TITLE OF INVENTION: THEIR USE  
; NUMBER OF SEQUENCES: 22  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Morrison & Foerster  
; STREET: 545 Middlefield Road, Suite 200  
; CITY: Menlo Park  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94025  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/640, 476  
; FILING DATE: 19910110  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kate H. Murashige  
; REGISTRATION NUMBER: 29,959  
; REFERENCE/DOCKET NUMBER: 24615-20009.20  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 327-7250  
; TELEFAX: (415) 327-2951  
; TELEX: 706141  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1164 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; HYPOTHETICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; ORGANISM: Streptomyces murinus  
; STRAIN: DSM 40091  
; FEATURE:  
;

NAME/KEY: CDS  
LOCATION: 1..1164  
IDENTIFICATION METHOD: experimental  
OTHER INFORMATION: /EC number=5.3.1.5  
OTHER INFORMATION: /product= "xylose isomerase (glucose isomerase)"  
OTHER INFORMATION: /evidence= EXPERIMENTAL  
OTHER INFORMATION: /standard\_name= "D-xylose ketol isomerase"  
US-07-640-476-6

Query Match 1.2%; Score 51.6; DB 1; Length 1164;  
Best Local Similarity 47.8%; Pred. No. 2.5e-05;  
Matches 150; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 31 AGCCCGCTCGAGAAATCATTAACCGGCTCAATACGACTACAACTGGGCTTCCAGTGT 90  
DB 16 ACCCCGAGAGACAGATTACCTTGGTGTGAGACCTGGCTGGCAGGAGGAGACCCG 75  
QY 91 GTGCGACACAACTCTACCCCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 150  
DB 76 TTGGGCGACGACCGCCCGCCCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 135  
QY 151 TTGGTGGCATGACAAATCTACAGAGCCCTGAATCTTCTACTGCGGAGAGATAC 210  
DB 136 CTGGGCGCTTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 195  
QY 211 TCCTGACACGAG 270  
DB 196 ACCGAGGCGAGTGCATACAGCGCTTCCGAGAGAGAGAGAGAGAGAGAGAG 255  
QY 271 GTGCGCAAG 330  
DB 256 GTGCGCAAG 315  
QY 331 GCTACTCGCGGCGCA 344  
DB 316 GCCAAGACGCGCGCA 329

RESULT 2  
US-08-811-583-1  
Sequence 1, Application US/08811583  
Patent No. 6218142

GENERAL INFORMATION:  
APPLICANT: Maassenegger, Michael  
APPLICANT: Riedel, Leonhard  
APPLICANT: Schiebel, Winfried  
APPLICANT: Sanger, Heinz  
TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING  
TITLE OF INVENTION: POLYPEPTIDES HAVING THE ENZYMATIC ACTIVITY OF AN  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: FISH & NEAVE  
STREET: 1251 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10020  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/811,583  
FILING DATE: 05-MAR-1997  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: Haley, James F.  
REGISTRATION NUMBER: 27,794  
REFERENCE/DOCKET NUMBER: MPG-1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-596-9000

TELEFAX: 212-596-9090  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3731 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
ORIGINAL SOURCE:  
ORGANISM: Tomato  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 194..3535  
US-08-811-583-1

Query Match 1.2%; Score 49; DB 3; Length 3731;  
Best Local Similarity 51.6%; Pred. No. 0.00041;  
Matches 112; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 2850 TACACTCTATCGAGCTGATGATCTCTGTCGGCGGATCCCAAGCCATTTCCCTAGTGA 2309  
DB 2425 TAATTTCATTCTGAGAGGAATGTGTTGGCAAAAATCCATGCTTGATCTGCTGATGA 2484  
QY 2910 TATCCAAAGGCTTCAGAGCTTCAAGCAGAGCTCCACAGTCTCAAGATGTAATCAT 2369  
DB 2485 TATTCGTGTTTAAAGCTGTAAATGTTTGAGGCTGACACACATGATGATGTTGT 2544  
QY 2970 CTTCCTACTAAGAGAGATGATACCGCTTGCTAAGAGATATCTGTGAGACTACAGACG 3029  
DB 2545 ATTCCCTCAGAAAAGAAAAGAACCTCATCCGATGATGATGTTCTGGAGATGATTTGGATGG 2604  
QY 3030 CGATATGCGCTGGGCTGCTGCGGATCCGAGATGCTC 3066  
DB 2605 GGATATCTACTTGTGTTGCTGGGATCAAGACATGATC 2641

RESULT 3  
US-08-986-304-1  
Sequence 1, Application US/08986304A  
Patent No. 6184343

GENERAL INFORMATION:  
APPLICANT: Stamatoyanopoulos, George  
APPLICANT: Papayannopoulou, Thalia  
APPLICANT: Yang, Yi  
TITLE OF INVENTION: FETAL GLOBIN INDUCING FACTOR  
FILE REFERENCE: 64657.0102  
CURRENT APPLICATION NUMBER: US/08/986,304A  
CURRENT FILING DATE: 1997-12-05  
EARLIER APPLICATION NUMBER: 60/033,247  
EARLIER FILING DATE: 1996-12-06  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 1  
LENGTH: 1318  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-08-986-304-1

Query Match 1.0%; Score 44; DB 3; Length 1318;  
Best Local Similarity 51.0%; Pred. No. 0.0078;  
Matches 104; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 2506 AATGCCCAAGTGAATTTGCGCAATGGGTTTACGAGAGTATTTCAAGTCGCGCACTCGA 2565  
DB 592 AATACCAAGAGTGTCTTTCTTACTAGCAGCATGATGACAGATATCAATGCCCAACAAA 651  
QY 2566 GTAGACGAGGCGCTGTGCTTTTCTTGTGGGCTACCTGACAGTCAAGAGAGACACTG 2625  
DB 652 GGCTCTTGAACCCCTTGCATCTTGCTGGGAGACAGACGACAGAGATACCTAGAA 711  
QY 2626 AACTCTTGATGAAGAGTGGGTTGATCCCAAGAGAGAGAGAGAGAGAGAGAGAGAG 2685  
DB 712 CTCCTCTGATGAACTTATGCTTAAACCGAGGCTGAAAAACAATTTGAGAGAACTGCA 771











```

; LENGTH: 36093
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: m1ec_feature
; LOCATION: (1)..(36093)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14664

```

```

Query Match      0.9%; Score 38.6; DB 4; Length 36093;
Best Local Similarity 52.9%; Pred. No. 5.3;
Matches 83; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

```

```

QY      24  GAGGATAGCCCGCTCGAGAAATCATTAACCGGCTCAATAGCACTTACAACTGGGCT 83
Db      16851 GATGAGACAGCTTATCAAGGTGACTGAGTGGCAGACATACACTTACGACTCGGTAT 16910

QY      84  CCACTGTGTGCGAGACACAACTCTCACCCCGCCGGAAGAGAGCTGGCCGAGAGTGA 143
Db      16911 CCACTCGGGGGCCAAACACTGCTGCTCCCTCCGTCAGCAGCAAGGCATCATGAGAGAGA 16970

QY      144 CGAGGATTTGGTGGCCATGACAAAGATTTACAGAGCC 180
Db      16971 TGAGGCTTGGGGCCAGTACAGCTCAAGAAAAACC 17007

```

Search completed: August 10, 2005, 13:52:33  
 Job time : 629 secs

**This Page Blank (uspto)**





APPLICATION NUMBER: US 08/811,583  
 FILING DATE: 05-MAR-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Haley, James F.  
 REGISTRATION NUMBER: 27,794  
 REFERENCE/DOCKET NUMBER: MPG-1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-596-9000  
 TELEFAX: 212-596-9090  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 3731 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 ORIGINAL SOURCE:  
 ORGANISM: Tomato  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 194..3535  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:

Query Match	1.2%	Score 49	DB 9	Length 3731
Best Local Similarity	51.6%	Pred. 0.00041		
Matches 112	Conservative	0	Mismatches 105	Indels 0
			Gaps	0

QY	2850	TACACTCCATACGACGCTGTAGTATGTCCTGTGTGGCGAGATCCCAAGCCATTTCCCTTAGGA	2908
QY <td>2850</td> <td>TAAATTTCACTTCGTGAAGGAAATGTGTGTGTGCAAAAATAATCATCTTCATCTCGGTGA</td> <td>2484</td>	2850	TAAATTTCACTTCGTGAAGGAAATGTGTGTGTGCAAAAATAATCATCTTCATCTCGGTGA	2484
QY <td>2910</td> <td>TATCCAAAGGGTTTCGACGAGTCTTCAAGCCAGAGCTCCACAGTCTCAAGATGTATCAT</td> <td>2968</td>	2910	TATCCAAAGGGTTTCGACGAGTCTTCAAGCCAGAGCTCCACAGTCTCAAGATGTATCAT	2968
QY <td>2485</td> <td>TATTCGTGTTTTAAAGGCTGTAAATGTTTCGAGCGCTGCACACATGTAGATTTGTTGT</td> <td>2544</td>	2485	TATTCGTGTTTTAAAGGCTGTAAATGTTTCGAGCGCTGCACACATGTAGATTTGTTGT	2544
QY <td>2970</td> <td>CTTCTCTTAATAAGAGATGTATCCGCTTCTTAAGAAAGCTATCTGTGTGAGACTAGACGG</td> <td>3029</td>	2970	CTTCTCTTAATAAGAGATGTATCCGCTTCTTAAGAAAGCTATCTGTGTGAGACTAGACGG	3029
QY <td>2545</td> <td>ATTCCCTCAGAAAGGAAAAAGACCTCATCCGAAATGATGTTCTGGAGATGTATTTGGATGG</td> <td>2604</td>	2545	ATTCCCTCAGAAAGGAAAAAGACCTCATCCGAAATGATGTTCTGGAGATGTATTTGGATGG	2604
QY <td>3030</td> <td>CGATATGGCCTTGGGTCTGTCTGGGATCCGAGATCGTC</td> <td>3066</td>	3030	CGATATGGCCTTGGGTCTGTCTGGGATCCGAGATCGTC	3066
QY <td>2605</td> <td>GGATATCTACTTTGTTGTCTGGGATCAGACATGATC</td> <td>2641</td>	2605	GGATATCTACTTTGTTGTCTGGGATCAGACATGATC	2641

```

US-RESULT 2
US-10-156-761-7161
? Sequence 7161, Application US/10156761
? Publication No. US20030119018A1
? GENERAL INFORMATION:
? APPLICANT: OMURA, SATOSHI
? APPLICANT: IKEDA, HARO
? APPLICANT: ISHIKAWA, JUN
? APPLICANT: HORIKAWA, HIROSHI
? APPLICANT: SHIDA, TADAYOSHI
? APPLICANT: SAKAKI, YOSHIYUKI
? APPLICANT: HATTORI, NASHAIIRA
? TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
? FILE REFERENCE: 249-262
? CURRENT APPLICATION NUMBER: US/10/156,761
? CURRENT FILING DATE: 2002-05-29
? PRIOR APPLICATION NUMBER: JP 2001-204089
? PRIOR FILING DATE: 2001-05-30
? PRIOR APPLICATION NUMBER: JP 2001-272697
? PRIOR FILING DATE: 2001-08-02
? NUMBER OF SEQ ID NOS: 15109
? SEQ ID NO 7161
? LENGTH: 1164
? TYPE: DNA
? ORGANISM: Streptomyces avermitilis
? FEATURE:
? NAME/KEY: CDS
? LOCATION: (1)..(1164)

```

US-10-156-761-7161	
Query Match	1.1%;
Best Local Similarity	46.9%;
Matches 146; Conservative	0; Mismatches 165; Indels 0; Gaps 0

QY	31	GGCCCCCGGAGGAATATCATMAACCGGGCTCATATAAGACTAACACTGTTGGGCTCCAGTGT	90
Db	16	ACCCCCGAGGACAGGTTTAACTTTGGCTGTGGACCCGTTCGGCTGGCAGAGGAAGGACCCG	75
QY	91	GTGCGAGACACAACTTTCACCCGCCACCGCCGAGAGAGCTGGCCGAGATGACGAGAT	150
Db	76	TTGCGCGAGCGCACCCCGCGCGCCCTGGACCCCGGTGAGAGCGTGCACACGCGCTTCCCGG	135
QY	151	TTGCGTGCACATGACAAAGATCTACAGAGCCCTGAACTTTCTCTACTGCGGGAAGATGAC	210
Db	136	CTCGGTGCCACGAGATGACCTTCCACGACACAGACTGATTCCTTGGGTTCCTCGGAC	195
QY	211	TTCCCTGAACCAAGGAGAACCACTTCTTCAATCGAGGCTCAAAAGCTGCGAGCTCGAATCGG	270
Db	196	ACCGAGCCGAGTGCACATCAAGCGCTTCGGGAGGCGCTCGAGCGCCACCGGACGTGGCC	255
QY	271	GTGCCCCAAGGCCACGCGGACCTTGACACGCTTCGTTGTCGAAGAACTTCCCGCGCC	330
Db	256	GTGCGGATGGGCACCAACCACTTCTCAACGACACCCCGTCTTCAGAGAGGGGGCGTTTACC	315
QY	331	GCTACTGCCGG	341
Db	316	GCCACGACCG	326

```

US-10-156-761-1
; RESULT 3
; US-10-156-761-1
; Sequence 1, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHISA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 1
; LENGTH: 9025608
; TYPE: DNA
; ORGANISM: Streptomyces avermitilis
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (4187715)
; OTHER INFORMATION: a, t, c, g, other or unknown
; US-10-156-761-1

```

Query Match	1.1%	Score 47	DB 15	Length 9025608
Best Local Similarity	46.9%	Pred. No.	0.17	Mismatches
Matches 146	Conservative 0	Mismatches	165	Indels 0
				Gaps 0

QY	31	AGCCCGCTCGAGGAATATCATTAACCGGCTCATATACATCTCMAAACCTGGGCTTCCATGT	90
QY	8551009	ACCCCGAGGACAGCTTACCTTGGGCTTGGACCTCTCGACTCGGACGGAGAGGACCCG	85510688
Db	91	GTGCGACAGCACTTCATCCCTCCACCGCGGAGAGCTGGCCGAGCTGACGAGAT	150
QY	8551069	TTGCGGACAGCGACATCCGCGCGCTTGGATCCGGGTGAGACGGTGCAGCGCTTCGCGGGG	85511288



QY 871 AATCAGACCTACCCAGAGTATTTGAGGCTCTCTCTTACAGCCAGAGAGAGCGG 930  
D 348 S.H...S...D.MTHC.MT.N.S.H...MN.NM.WY.BD..SNSD.Y.BM..A. 289  
QY 931 CCTGTGATGCCAGCGCTTTTGAGGCTGACACTTGATGTTCTCTAGCAAGAGAGA 990  
D 288 TH..TNB.....SNMG.TG...TNRGY.GNCS.H.....N.NN.HH...MT.KNNR.R.. 229  
QY 991 ACAACCAAGTCCACATAGATAC 1014  
D 228 A.AS..BA.BTABNNHM.S.BM.M 205

RESULT 6  
US-10-146-731-358/C

Sequence 358, Application US/10146731  
Publication No. US20030129692A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P330R1C123  
CURRENT APPLICATION NUMBER: US/10/146,731  
CURRENT FILING DATE: 2002-05-15  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 358  
LENGTH: 1049  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-146-731-358

Query Match 1.1%; Score 44.8; DB 15; Length 1049;  
Best Local Similarity 5.2%; Pred. No. 0.0049;  
Matches 42; Conservative 229; Mismatches 531; Indels 2; Gaps 1;

QY 211 TCCTGGAACCGAGGAGCACTTCTTCATCGACCCAAAGCTGGAGCTGCACTGG 270  
D 1006 YMY...MSM...M.....BHEM.HK.Y.DS.R.T.M.N.AMTRHAKB.BAMMS... 947  
QY 271 GTGGCCAAAGCCGAGCCGACCTGACACGCTTCGTCGTCAGAGCACTCCCGGCGCC 330  
D 946 S.S.N..B..C.MHY..G.N.DM.Y.H.MTB..T.BW.ABT.HNAHRE.THRGH. 887  
QY 331 GCTACTCGCGGCAACATGGGCTTCAGACTGTGTGCTCGAGGTGCTTAATAGGTTT 390  
D 886 S..Y.RCM.MTMG.DR.RMBHM.R.DSTAKBK...SBS.S.S...NA.H..GAR.H.S 827  
QY 391 ATGGCACTCCCATTAACACACCGAGTGAACGTTTGGCAGAACTCTAAGGCGCCCAAGT 450  
D 826 .BS.CMTJC.CBGBHAT.R..AB.ADNBWB.WBTGAG..YND..KM.NN.BN... 767  
QY 451 GGCCTGAGCGCCGCACTTACCAACGCAAGAGTAGAGCCGCGCAATGTCACT 510  
D 766 SAM..K..MNS.H.RY...TH...RMA.SY..NM...NM--SYGNS.Y..BAA..N 709

QY 511 TTGCGATCCGCCAAACGCTGTGACTCGCTTCCACAGCTCTCTATTCAGGC 570  
D 708 DS.H.A..NM.G..MMWS..SM.CNMT.S.NM.N..KCH.BCS...S.SNMS.H.....M.. 649  
QY 571 GCGGCATACCGCTTAAGTTCCTCCGATCCAGTGAATACCGGTTCACCAAGACCATCTTC 630  
D 648 NM...RYNHC.YW.BH.DNCY...AY.S.S.KAYSASS.NHNHMK.M..BM.NMA.N. 589  
QY 631 GAGATGAGATCTCAATCATGCTGACCAAGCGGCGCAAGCTGTCTGATTAATGTT 690  
D 588 KDA.C.S..RDSMS.H.B..MD...TASD.H.YNNS.H.RY..T....SCN.A.S. 529  
QY 691 GCGGTCGCGCGCCCGCCCGCTTATGAGCGCTTTGACAGAGTACCACTGA 750  
D 528 .NCS.N.GM..S.D..HSSWB...SNMS.H.A.CRMRS.NBSK.ST.MNM.YGYTRMH 469  
QY 751 AGGCATGCCAATACGAGAGTCCACGCGCAGCATCATAGACGCGGACCGAGTGAT 810  
D 468 RY.RD...B...RS.BSAYTNSG.CB.SSHCS.S.MNBS.H.BM.YM..M.KS.MT.M. 409  
QY 811 TCCTTGATATCTCAAGGACCTCTATGTTGAGTGTCTTCAAGCGCTGCGTCA 870  
D 408 .MNC.H.B..N..ND..S.N.SM..M.BRCY.Y..M.SM.SS.T.S.NKSTYB...N. 349  
QY 871 AATCAGACCTACCCAGAGTATTTGAGGCTCTCTCTTACAGCCAGAGAGAGCGG 930  
D 348 S.H...S...D.MTHC.MT.N.S.H...MN.NM.WY.BD..SNSD.Y.BM..A. 289  
QY 931 CCTGTGATGCCAGCGCTTTTGAGGCTGACACTTGATGATGATCTCTAGCAAGAGAGA 990  
D 288 TH..TNB.....SNMG.TG...TNRGY.GNCS.H.....N.NN.HH...MT.KNNR.R.. 229  
QY 991 ACAACCAAGTCCACATAGATAC 1014  
D 228 A.AS..BA.BTABNNHM.S.BM.M 205

## RESULT 7

US-10-140-472-358/C  
Sequence 358, Application US/10140472  
Publication No. US20030138888A1  
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: Desnoyers, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gerritsen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William  
APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE REFERENCE: P330R1C168  
CURRENT APPLICATION NUMBER: US/10/140,472  
CURRENT FILING DATE: 2002-05-06  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 358  
LENGTH: 1049  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-140-472-358

Query Match 1.1%; Score 44.8; DB 15; Length 1049;

[illegible]

RESULT 8  
 US-10-141-761-358/c  
 ; Sequence 358, Application US/10141761  
 ; Publication No. US20030148432a1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Bersetti, Maureen  
 ; APPLICANT: DeForge, Laura  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerritsen, Mary E.  
 ; APPLICANT: Goddard, Audrey

```

APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C198
CURRENT APPLICATION NUMBER: US/10/141,761
PRIORITY FILING DATE: 2002-05-08
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 358
LENGTH: 1049
TYPE: PRT
ORGANISM: Homo Sapien
US-10-141-761-358

```

Query Match 1.1%; Score 44.8; DB 15; Length 1049;  
Best Local Similarity 5.2%; Pred.No. 0.0049;  
Matches 42; Conservative 229; Mismatches 531; Indels 2; Gaps 1;

Qy	211	TCGCTGAACAGGACGAGAGGCAACTTCTTATGATGAGGCCAAAGCTGAGCTGCAACTGCG	270
Dy	1006	YMY...MSM...M.....BBBM..HK.Y.DS.R.T.M.N.AMTRHAB.BAMS...	947
Qy	271	GTGCCCCAAGCCGACGCGGCACTGACACGCTTCGCTGTCAGAGAACTCCCGCGCC	330
Dy	946	S.S.N..B..C..WHY...G.N.DM.Y.H..MTB..T.BW.ABT..HMAHNB..THRGH..	887
Qy	331	GCTACTGCGGCGCAACAATGGGATTCGACAGCTGTCTGTCAGAGGCTTAATAGTTT	390
Dy	886	S..Y.RCM.MTWG.DR.RBHW.R.DSTAKBK...SBS.S.S...NA.H..GAR.H.S.827	
Qy	391	ATGCACCTCCCATATAACACACGAGCGGCAAGCTTTGGCGAAGTCTTAAGCGGCCCAAT	450
Dy	826	.BS.CMDTC.CBGAHHHT..R..AB.ADNBWB..WBTHGAG..YND..KM.NN.BN...	767
Qy	451	GGCCTGAGCGGCCCAACTCTTACCAACACCAAGGATGAGCGCCCAATGTCAT	510
Dy	766	SAM..K..MNS.H.RV...TH...RVA.SY..NNM...NM--.SVSGNS.Y..BA..N.709	
Qy	511	TTGCTGATCGGCCCAAGCGCTGTGATCTGCTTGCACAGAGTCTCTTATTCAGCG	570
Dy	708	DS.H.A..NM.G...MMWS..SM.CNMT.S.NM.N..KCH.BCS...S.SNNS.H...M..649	
Qy	571	GCGGCGATTACCCCTAAAGTTGCCCGGATTCAGTAAATACCGGTTCCAAGCAGCATCTTC	630
Dy	648	NM...RYNHC.YW.BH.DWCY...AY.S.S.KAYSHSS.HNNHK.M..BM.NMA.N.589	
Qy	631	GAGAGTGAATCTCAATCAATGCACACCAAGCGGCCCAAGGCGCAAGCTGTCTGATTAATGTT	690
Dy	588	KDA.C.S..RDSNS.H.B..MD...TASD..H.YNNS.H.RV..T....SCN.A.S.529	
Qy	691	GCGCTGCGCGCGCGCGCGCGCGCTGATTCGAGCGCTTTGCAAGGTAACGAGTCTCA	750
Dy	528	.NCS.N.GM..S.D..HSSMB...SNNS.H.A.CRRMG.S.NBSK.ST.NMM.YGSYTRM	469
Qy	751	AGGATGCCAATAGCAGAGATCCGACGCGGACAGATCATGAGAGCGAGCGACAGTGCAT	810
Dy	468	RY.RD...B..RS.BSATNSG.CB.SHKC.S.MNBS.H.BM.YM..M.KS.NT.M..409	
Qy	811	TCCTTGATACATCTCAAGGCACTTCTTAATGTTGAGAGTCTTCAAGCTTCGCTGAC	870
Dy	408	.NAC.H.B..N..ND..S.N.SM..M.BRCY.Y..M.SM.SS.T.S.NKSTYB...N.349	
Qy	871	AATCAAGACATACCAAGAGTAGTATTGAGGCTCTCTTACAGCCCAAGAGAAAGCG	930
Dy	348	S.H...S...D..MTHC..MT..N.S.H...MN.NM.WY..BD..SNSD.Y.BM..A.289	

QY 931 CCTGTGATGCCACGGTCTTTGAGCTGGAACATTGATGCTCCTTACCAAGAGA 990  
D 288 TH.TNB.....SNMNG.TG...TNRGY.GNCS.H.....N.NN.HH...MT.KNNR.R... 229  
QY 991 ACAACCAAGTCCCAATAGATTAAC 1014  
D 228 A.AS..BA.BTABNNHM.S.BM.M 205

## RESULT 9

US-10-142-885-358/C

; Sequence 358, Application US/10142885  
; Publication No. US20030157604A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerlitsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P333081C248  
; CURRENT APPLICATION NUMBER: US/10/142,885  
; CURRENT FILING DATE: 2002-05-10  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 358  
; LENGTH: 1049  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-142-885-358

Query Match 1.1%; Score 44.8; DB 16; Length 1049;

Best Local Similarity 5.2%; Pred. No. 0.0049;  
Matches 42; Conservative 229; Mismatches 531; Indels 2; Gaps 1;

QY 211 TCCTGACCAAGCAAGCAACTTCTTATCGAGGCCAAAGCTGCGAGCTGCACTGG 270  
D 1006 YMY...MSM...M.....BHBH.HK.Y.DS.R.T.M.N.AMTRMHAKB.BAMMS... 947  
QY 271 GTGGCCAAAGCCGCGGCGGACCCGCGACACGCTTCCGCTGTCGAAGAACTCCCGCGCC 330  
D 946 S.S.N..B..C...WHY..G.N.DM.Y.H..MTB..T.BW.ABT.HVAHNB..THRGH. 887  
QY 331 GCTACTCCGCGCAACATGGGACATGCTGTGTCGAGTGTCTTAATAGGTTT 390  
D 886 S..Y.RCM.MTMG.DR.RWBHW.R.DSTAKBK...SBS.S.S...NA.H.GAR.H.S 827  
QY 391 ATGCCACCTCCCAATTAACACACGAGTGAAGCTTTGGCAGAACTTAAGCGGCCAAGT 450  
D 826 .BS.CMDTC.CBGBAHAT.R..AB.ADNBWB.WBTHGAG..YVDD...KM.NN.BN... 767  
QY 451 GGCGTAGCGCGCCCAACTCTACCAACAAAGCAAGATGAGCCCGCCCATGTACT 510  
D 766 SAM..K..MNS.H.RY...TH...RMA.ST..NM..NM--..SYGNS.Y..BAA..N 709  
QY 511 TTGCGTATCCGCAAGCTGCTGATGCTGCTGCTGCAAGGCTCCTCATTCACGGC 570  
D 708 DS.H.A..NM.G..MMS.SM.CNMT.S.NM.N..KCH.BCS...S.SNMS.H....M. 649

QY 571 GCGCGATACCCCTTAAGTTCGCCGATCCAGTGAATACCGGTTCCAAAGCATTCTTC 630  
D 648 NM...RYNHC.YW.BH.DNCY...AY.S.S.KAYSASSS.HNHNK.M..BM.NNA.N. 589  
QY 631 GAGATGAGATTCATCATGATGTCACCAAGCGGCCAAGGCGCAAGCTGTGTAATGTT 690  
D 588 KDA.C.S..RDSNS.H.B..MD...TASD..H.YNNS.H.RY..T....SCN.A.S. 529  
QY 691 GCGCTGCGCGCGCGCGCGCGCGCTTATGCGAGCGCTTTGACAAGATACCACTCGA 750  
D 528 .NCS.N.GM..S.D..HSSMB...SNMS.H.A.CMRGS.NBSK.ST.NMM.YGYTRMH 469  
QY 751 AGCATGCCAATTCAGAGATCCACGCGACAGGTGATAGACGACGCGAGTGGAT 810  
D 468 RY.RD...B...RS.BSAYTNG.CB.SSHCS.S.MNS.H.BM.YM..M.XS..NT.M. 409  
QY 811 TCCTTGATCATCTCAAGCACTTCTTATGTTGAGTGTCTTCAAGCGTCCGTCAC 870  
D 408 .NAC.H.B..N..ND..S.N.SM..M.BRCY.Y..M.SM.SS.T.S.NKSTYRB...N. 349  
QY 871 AATCAGACACTACCCAGAGTATTGAGGCTCTCTTACAGGCCAGAGAAAGCG 930  
D 348 S.H...S...D..MTHC.MT..N.S.H...MN.NM.WY..BD..SNSD.Y.BM..A. 289  
QY 931 CCTGTGATGCCACGGTCTTTGAGCTGGAACATTGATGCTCCTTACCAAGAGA 990  
D 288 TH.TNB.....SNMNG.TG...TNRGY.GNCS.H.....N.NN.HH...MT.KNNR.R... 229  
QY 991 ACAACCAAGTCCCAATAGATTAAC 1014  
D 228 A.AS..BA.BTABNNHM.S.BM.M 205

## RESULT 10

US-10-158-790-358/C

; Sequence 358, Application US/10158790  
; Publication No. US20030180879A1  
; GENERAL INFORMATION:  
; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: DeForge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Gerlitsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P333081C248  
; CURRENT APPLICATION NUMBER: US/10/158,790  
; CURRENT FILING DATE: 2002-05-10  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 358  
; LENGTH: 1049  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-158-790-358

Query Match 1.1%; Score 44.8; DB 16; Length 1049;

Best Local Similarity 5.2%; Pred. No. 0.0049;  
Matches 42; Conservative 229; Mismatches 531; Indels 2; Gaps 1;

QY 211 TCCTGACCAAGCAAGCAACTTCTTATCGAGGCCAAAGCTGCGAGCTGCAACTGG 270

Db 1006 YMY...MSM...M.....BHEM.HK.Y.DS.R.T.M.N.AMTRHNAAB.BAMMS... 947

Qy 271 GTGCCCAAGCCCGCCGACCCTTGACAGCGCTTCGGTGGTCCAAGAACTCCCGCGCC 330

Db 946 S.S.N..B..C..WHY...G.N.DM.Y.H..MTB..T.BM.ABT.HMAHNB..THRGH.. 887

Qy 331 GCTACTCCCGGCCCAACATGGCGCATTTGCAGACTGTGTGCTGAGGTGCTTAATAGGTTT 390

Db 886 S..Y..RCM.MTNG..DR..RWBHW..R..DSTAKKX...SBS.S.S....NA.H..GAR.H.S 827

Qy 391 ATGCCACCTCCCAATATACACACCGATCCAGCTTTGGCAGACTCTAAGCGGCCCACT 450

Db 826 .BS..CMDTC..CGABHAT..R..AB..ADNMBWB..WBTHGAG..YND..K.M.NN..BN... 767

Qy 451 GGCCTGAGCGCCGCCAACCTCTACACACCAACCAAGAGATGAGCGCCGCAATGTCACT 510

Db 766 SAM..K..MNSS.H.RY..TH...RMA..SY..NNM...NM--.SYSGNS.Y..BAA..N 709

Qy 511 TTGCGTGATCGGCCCAAGCGTGTGTTGACTGCTGTCCACAGACTCTCTATTACGGC 570

Db 708 DS.H.A..NM.G...MMMS..SM.CNMT.S.NM.N..KCH.BCS...S.SNMS.H....M.. 649

Qy 571 GCGGCGATACCCCTTAAGTTTCCCCGATCCAGTAATAACCGCTTCCAAGCACATCTCTC 630

Db 648 NM...RVNHC.YW..BH..DNCY...AY.S.S..KAYSASSS.NHNHKK.M..BM.MNA.N.. 589

Qy 631 GAGAGTGAGAAATCTCAATCAGTGACACCAAGCGGCCAAGGCGAAGCTGTGTATAATGTT 690

Db 588 KDA.C.S..RDSNNS.H.B..MD...TMSD..H.YNNS.H.RY..T....SCN.A.S.. 529

Qy 691 GCCGCTGCCGCCGCCGCCGCCGCTGCTATTTGCGAGCGCTTTGACAAGTATGACACTGA 750

Db 528 .NCS.N.GM..S.D..HSSMB...SNMS.H.A.CRRKGS.NBSK.ST.NNM.YGTYRMH 469

Qy 751 AGCGCATGCCAATACGAGAGATCTCCACGCGGACAGCGTCATATAGACGAGCGGACGAGTGAT 810

Db 468 RY.RD...B...RS..BSAYTNSG..CB..SSHCS.S..MNSB.H..BM.YM..M.KS..MT..M.. 409

Qy 811 TCCTTTGATATCATCTCGAAGCACTTCTCTATGTTGAGTGTCTTTCAGCGCTCCCGTCCAC 870

Db 408 .NAC.H.B..N..ND..S.N.SM..M.BRCY.Y..M.SM.SS..T.S..NKSITYB...N.. 349

Qy 871 AATCAGAGCACTACCCAGAGTAGTATTTTGAGAGCTCTCTCTTACACAGCCCAAGAGAAGCG 930

Db 348 S.H....S..D..MTHC..MT..N.S.H...MN..NM.WY..BD..SNBD.Y.BM..A.. 289

Qy 931 CCTGTGATGCCACGCTCTTTGAGGCTGTGACCTTGACATTGATGAGTCTCTTACGAAGAAGA 990

Db 288 TH.TNB....SNMNG..TG...TNRGY..GNCS.H....N.NN.HH...MT..KNNR.R... 229

Qy 991 ACAACCAAGTCCCACTAGATAAC 1014

Db 228 A.AS..BA..BTABNNHM.S.BM..M 205

RESULT 11  
US-10-137-871-358/c  
; Sequence 358, Application US/10137871  
; Publication No. US20030207350A1

APPLICANT:	Baker, Kevin P.
APPLICANT:	Bersini, Maureen
APPLICANT:	DeFogers, Luc
APPLICANT:	Demoyers, Laura
APPLICANT:	Filyaroff, Ellen
APPLICANT:	Gao, Wei-Qiang
APPLICANT:	Gertelsen, Mary E.
APPLICANT:	Goddard, Audrey
APPLICANT:	Godowski, Paul J.
APPLICANT:	Gurney, Austin L.
APPLICANT:	Shemwood, Steven
APPLICANT:	Smith, Victoria

```

APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: ACIDS ENCODING THE SAME
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C153
CURRENT APPLICATION NUMBER: US/10/137,871
CURRENT FILING DATE: 2002-05-03
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 358
LENGTH: 1049
TYPE: PRT
ORGANISM: Homo Sapien
US-10-137-871-358

```

Query Match 1.1%; Score 44.8; DB 17; Length 1049;  
Best Local Similarity 5.2%; Pred. No. 0.0049;  
Matches 42; Conservative 229; Mismatches 531; Indels 2; Gaps 1;

OY		211	TCCCTGAACCGAGGAGAAGCAACTTTCATTCGAGGCCAAAGCTGGCAGCTGCACCTG	270
Db	:::	1006	YMY...MSM...M.....BHM.HK.Y.DS.R.T.M.N.AMTRMHAK.BAMWS...	947
OY		271	GTCGCCAAGCGCACGCGACCCTTGACACGCTTCGCTGTGCACAGAAACCTCCCGCGCC	330
Dd	:	946	S.S.N..B..C..MHY..G.N.DM.Y.H..MTB..T..BW..ABT..HNHNR..THRGH..	887
OY		331	GCTACTGCGCGCCAACAATGGCATTTGCAGACTGTGTGTTGCTCGAGTGCCTTAATAGTTT	390
Dd	:	886	S..Y..RCM.MTMG..DR..RWBHM..R.DSTAKRK...SBS..S.S...NA..H..GAR..H.S	827
OY		391	ATGCCACTCCCAATAACACACCGGTGCAAGCTTTGGCAGAACTCTAAGGGCCCCAAGT	450
Dd	:	826	.BS.CMDTC.CVBGBHAT..R..AB..ANBWB..WBTGAG..YND..XV..NN..BN...	767
OY		451	GGCCTGAGCGCGCCAACCTTACCAACACCAAACGCAAGATGAGCGCGCAATGTCTACT	510
Dd	:	766	SAM..K..MNSS.H.RY...TH...KDA.ST..NMN...NM--SYSNS.Y..BA..N	709
OY		511	TTGCTGATCCGCCCAACAGCTCGTTGATCTGCTGTGCAACAGTCTCTATTACCGC	570
Dd	:	708	DS.H.A..NM.G..MMMS.SM.CMTT.S.NM.N..KCH.BCS...S.SNMS.H...M.	649
OY		571	GCGCGATACCCCTPAAAGTTCCCGCATCCAGTAGTAATCCGTTCCAAACGACATCTCTC	630
Dd	:	648	NM...RYNH.K.YW..BH.DNCY...AY.S.S.KAYSASS..HNHNKK..M..BM..NMA..N.	589
OY		631	GAGAGTAGAATCTCAATCAGTAGTGACCAAGCGGCGCCAAGGCAAGCTGTGTGATTAATGT	690
Dd	:	588	KDA..C.S..RDSNRS.H.B..MD...TASD..H.YNNS..H.RY..T....SCN..A.S.	529
OY		691	GCCGCTGCGCGCGCGCGCGCGCGCTTCATTGTCGAGCGCTTTGGACAAGTAGACGACTCGA	750
Dd	:	528	.NCG.N.GM..S.D..HSSMB...SNMS..H.A.CMRGS..NBKS.ST.MMN.YESYTRMH	469
OY		751	AGGCAATGCATATACGAGAGATCCACGCGCACAGAGTATAGACGAGCGACAGAGTGAT	810
Dd	:	468	RY..RD...B...RS.BSAYTNSG.CB..SSHCS..S.MNBS..H..BM..YM..M.KS..NT..M.	409
OY		811	TGCTTTGATATCATCGAACGACCTTCCTATGAGTGTCTGAGTGCTTCAGGCGCTTGGCGCTAC	870
Dd	:	408	.MAC.H.B..N..ND..S.N.SM..M..BRCY..Y..M..SM..SS..T.S..NKSTRB...N.	349
OY		871	AATCAGAGCACTACCAAGAGTAGTTTGGAGCTCTCTTCACAGCCACGAGAGAAGCGG	930
Dd	:	348	S.H...S...D..MTHC..MT..N.S.H...MN..NM.WY..BD..SNSD.Y..BM..A.	289
OY		931	CCTGTGATGCCACGCTTTTGGAGCTGAGACATTTAGTAGTCTCTCAGCAAGAAGAGA	990
Dd	:	288	TH..TNB...SNMG..TG...TURGY.GNCS..H...N..NN..HH...MT..KNR..R..	229

QY 991 ACAACCAAGTCCCATAGATAAC 1014  
| : : : : :  
Db 228 A.AS..BA.BTANNHM.S.BM.M 205

## RESULT 12

US-10-140-923-358/c  
; Sequence 358, Application US/10140923  
; Publication No. US20030207355A1  
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Geritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P330R1C18  
; CURRENT APPLICATION NUMBER: US/10/140, 923  
; CURRENT FILING DATE: 2002-05-07  
; Prior Application removed - See Palm or File Wrapper  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 358  
; LENGTH: 1049  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-140-923-358

Query Match 1.1%; Score 44.8; DB 17; Length 1049;

Best Local Similarity 5.2%; Pred. No. 0.0049;  
Matches 42; Conservative 229; Mismatches 531; Indels 2; Gaps 1;

QY 211 TCCCTGACCGAGCAAGCACTTCTTCATCGAGCCAAAGCTGCGAGCTGCACTGG 270  
| : : : : :  
Db 1006 YMY...MSM..M.....BHEM.HK.Y.DS.R.T.M.N.AMTNMAKB.BAMMS... 947  
QY 271 GTGCCCCAAGCCGAGCGGAGCTGAGACGCTTCGTTGTCAGAGAACCTCCCGCGCC 330  
| : : : : :  
Db 946 S.S.N..B..C..WHY...G.N.DM.Y.H..MTB..T.BW.ABT.HMAHRB..THRGSH. 887  
QY 331 GCTACTGCGCGGCAACATGAGCTTGCAGACTGTGTTGTCGAGGTCCTTAAGTTT 390  
| : : : : :  
Db 886 S..Y.RCM.MTMG..DR.RWBHM.R.DSTAKBK...SBS.S.S...NA.H..GAR.H.S 827  
QY 391 ATGCCACTTCCCATATAACACACAGTGCAGCTTTGCGAAGTCTTAAGGCGCCCAAGT 450  
| : : : : :  
Db 826 .BS.CMTTC.CBGBHAT.R..AB.ADNBWB.BHTGAG..YVDD...KM.NN.BN... 767  
QY 451 GGCCTGAGCGCGCACTTACCAACACAAAGCAAGATGAGCCCGCAATGTCACT 510  
| : : : : :  
Db 766 SAM..K..MNS.S.H.RV...TH...RMA.SY...NNM...NM--..SYGNS.Y..BAA..N 709  
QY 511 TTGCTGATCGGCCCAAGCTGTTGACTGCTCTGCGACAGTCTCTCTATTCACGGC 570  
| : : : : :  
Db 708 DS.H.A..NM.G...MMS..SM.CMTT.S.NM.N..KCH.BCS...S.SNMS.H...M.. 649  
QY 571 GCGGCAATACCGCTAAGTTCCCGATCGAGTGAATGAGTTCCAAAGCACTCTCTC 630  
| : : : : :  
Db 648 NM....RYNHC.YW.BH.DNCY...AY.S.S.KAYSASS.S.HNNKR.K..BM.NMA.N. 589

QY 631 GAGAGTGAATCTCATCTAGTGCACCAAGCGGCGCAAGGCAAGCTGTCTGATAATGTT 690  
| : : : : :  
Db 588 KDA.C.S..RDSNS.H.B..MD....TASD..H.YNNS.H.RY..T....SCN.A.S. 529  
QY 691 GCGCTGCG 750  
| : : : : :  
Db 528 .NCS.N.GM..S.D..HSSMB...SNMS.H.A.CRMGRS..NBSK..ST.MNM.YGSYTRM 469  
QY 751 AGCATGCCAATAGAGATGCCAGAGTCCAGAGTCAATAGACAGCGACGAGTGAT 810  
| : : : : :  
Db 468 RY.RD....B...RS.BSATYNG.CB.SSHCS.S.MNBS.H.BM.YM..M.KS.NT.M. 409  
QY 811 TCTTTGATACATCTCAAGCACTTCTCTATGTTGTCAGTGTCTTACGCGTCCGTCAC 870  
| : : : : :  
Db 408 .NAC.H.B..N..ND..S.N.SM..M.BRCY.Y..M.SM.SS.T.S.NKSTYRB...N. 349  
QY 871 AATCAGAGCACTACCGAGTACTTTTGAGGCTCTCTTCAAGCCCGAGAGAGCGG 930  
| : : : : :  
Db 348 S.H...S...D..MTHC..MT..N.S.H...MN.NM.WY..BD..SNSD.Y.BM..A. 289  
QY 931 CCTGTGATGCCAGGCTTTTGAGGCTGACACTTGATTGAGTCTCTGACAAAGAGA 990  
| : : : : :  
Db 288 TH.TNB....SNMG.TG...TNRGY.GNCS.H....N.NN.HH...MT.KNNR.R.. 229  
QY 991 ACAACCAAGTCCCATAGATAAC 1014  
| : : : : :  
Db 228 A.AS..BA.BTANNHM.S.BM.M 205

## RESULT 13

US-10-141-756-358/c  
; Sequence 358, Application US/10141756  
; Publication No. US20030207355A1  
; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.  
; APPLICANT: Beresini, Maureen  
; APPLICANT: Deforge, Laura  
; APPLICANT: Desnoyers, Luc  
; APPLICANT: Filvaroff, Ellen  
; APPLICANT: Gao, Wei-Qiang  
; APPLICANT: Geritsen, Mary E.  
; APPLICANT: Goddard, Audrey  
; APPLICANT: Godowski, Paul J.  
; APPLICANT: Gurney, Austin L.  
; APPLICANT: Sherwood, Steven  
; APPLICANT: Smith, Victoria  
; APPLICANT: Stewart, Timothy A.  
; APPLICANT: Tumas, Daniel  
; APPLICANT: Watanabe, Colin K  
; APPLICANT: Wood, William  
; APPLICANT: Zhang, Zemin  
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
; FILE REFERENCE: P330R1C200  
; CURRENT APPLICATION NUMBER: US/10/141, 756  
; CURRENT FILING DATE: 2002-05-08  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 550  
; SEQ ID NO 358  
; LENGTH: 1049  
; TYPE: PRT  
; ORGANISM: Homo Sapien  
US-10-141-756-358

Query Match 1.1%; Score 44.8; DB 17; Length 1049;

Best Local Similarity 5.2%; Pred. No. 0.0049;  
Matches 42; Conservative 229; Mismatches 531; Indels 2; Gaps 1;

QY 211 TCCCTGACCGAGCAAGCACTTCTTCATCGAGCCAAAGCTGCGAGCTGCACTGG 270  
| : : : : :  
Db 1006 YMY...MSM..M.....BHEM.HK.Y.DS.R.T.M.N.AMTNMAKB.BAMMS... 947  
QY 271 GTGCCCCAAGCCGAGCGGAGCTTGCAGACTGTGTTGTCGAGGTCCTTAAGTTT 390  
| : : : : :  
Db 946 S.S.N..B..C..WHY...G.N.DM.Y.H..MTB..T.BW.ABT.HMAHRB..THRGSH. 887



Db 946 S.S.N..B..C..WHY...G.N.DM.Y.H..MTB..T.BW.ABT.HMAHNB..THRGH.. 887  
Qy 331 GCTACTGCGGCGCAACATGGGCGATTTGCACTGTGTTGCTCGAGTGTCTTAATAGTTT 390  
Db 886 S..Y.RCM.MTMG..DR.RWBHM.R.DSTAKKBK...SBS.S.S...NA.H..GAR.H.S 827  
Qy 391 ATGCACACTCCCAATTAACACACAGGATGTTGGACAGACTTAAGGGCGCAAGT 450  
Db 826 .BS.CMDTC.CBGBBHT.R..AB.ADNBWB.BETHAG..YND..KM.NN.BN... 767  
Qy 451 GGCCTGAGCCGCCCACTTACCAACCAACGAGATGAGCCGCCCAATGTCACT 510  
Db 766 SAM..K..MNS.S.H.RY...TH...RMA.SY..NNM...NM--..SYGNS.Y..BAA..N 709  
Qy 511 TTGCGTATCGGCGCAACGCTGTTGACTGCTGTGCGACAGTCTCTTAATCAGGC 570  
Db 708 DS.H.A..NM.G..MMWS.SM.CNMT..S.NM.N..KCH.BCS...S.SNMS.H...M.. 649  
Qy 571 GCGGCGATACCCCTAAAGTTCCCGGATCCAGATACCGGTTCCAAACGACATCTC 630  
Db 648 NM...RYNHC.YW.BH.DNCT...AY.S.S.KAYSASS.HNHMK..M..BM.NMA.N. 589  
Qy 631 GAGAGTGAATCTCAATCACTGACGACCAAGCGGCGCAAGCTGTGATATGTT 690  
Db 588 KDA.C.S..RDSNS.H.B..MD...TASD..H.YNNS.H.RY..T....SCN.A.S. 529  
Qy 691 GCGCGTCCCGCGCGCGCGCGCTGCTTATGGCAGCGCTTTGACAGATCCGACTGA 750  
Db 528 .NCS.N.GM..S.D..HSSMB...SNMS.H.A.CMRGS.NBSK..ST.MNM.YGSYTRMH 469  
Qy 751 AGGATGCAATAGAGAGATCCCGCGGACAGTCAATAGACGCGGACCGAGTGTAT 810  
Db 468 RY.RD...B...RS.BSAYTNSG..CB.SSHCS.S.MMS.H.BM.YM..M.KS.NT.M. 409  
Qy 811 TCCTTGATATCATCTCAAGCACTTCTATGTTGATGTTGATGAGCTTGCGCTGAC 870  
Db 408 .NAC.H.B..N..ND..S.N.SM..M.BCY.Y..M.SM.SS.T.S.NKSTYRB...N. 349  
Qy 871 AATCAGAGCACTACCCAGAGTAGTTTGAAGCTCTCTTCAACAGCCAGAGAGCGG 930  
Db 348 S.H...S...D..MTHC..MT..N.S.H...NM.NM.WY..BD..SNSD.Y.BM..A. 289  
Qy 931 CCTGAGATGCGACGCTTTTGAAGCTGACACTTGATGAGTCTCTTCAAGAGAGA 990  
Db 288 TH.TNB...SNMG..TG...TNKGY.GNCS.H....N.NN.HH...MT.KNNR.R.. 229  
Qy 991 ACAACCAAGTCCACATAGATAC 1014  
Db 228 A.AS..BA.BTABNNHM.S.BM.M 205

RESULT 14  
US-10-141-759-358/c  
Sequence 358, Application US/10141759  
Publication No. US20030207361A1  
GENERAL INFORMATION:  
APPLICANT: Baker, Kevin P.  
APPLICANT: Beresini, Maureen  
APPLICANT: DeForge, Laura  
APPLICANT: DeNovo, Luc  
APPLICANT: Filvaroff, Ellen  
APPLICANT: Gao, Wei-Qiang  
APPLICANT: Gettleisen, Mary E.  
APPLICANT: Goddard, Audrey  
APPLICANT: Godowski, Paul J.  
APPLICANT: Gurney, Austin L.  
APPLICANT: Sherwood, Steven  
APPLICANT: Smith, Victoria  
APPLICANT: Stewart, Timothy A.  
APPLICANT: Tumas, Daniel  
APPLICANT: Watanabe, Colin K  
APPLICANT: Wood, William

APPLICANT: Zhang, Zemin  
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
FILE OF INVENTION: ACIDS ENCODING THE SAME  
FILE REFERENCE: P33081C197  
CURRENT APPLICATION NUMBER: US/10/141.759  
CURRENT FILING DATE: 2002-05-08  
Prior Application removed - See File Wrapper or Palm  
NUMBER OF SEQ ID NOS: 550  
SEQ ID NO 358  
LENGTH: 1049  
TYPE: PRT  
ORGANISM: Homo Sapien  
US-10-141-759-358

Query Match 1.1%, Score 44.8; DB 17; Length 1049;  
Best Local Similarity 5.2%; Pred.No.0.0049;  
Matches 42; Conservative 229; Mismatches 531; Indels 2; Gaps 1;

Qy 211 TCCCTGAACCGAGAGCAAGCACTTCTTCAATCAGGCGCAAGAGCTGCGAGCTCAACTG 270  
Db 1006 YMY...MSM...M.....BHM.HK.Y.DS.R.T.M.N.ATTHHAKB.BAMMS... 947  
Qy 271 GTGCCAAAGCCGCGGACCGCTGACAGCTTCTGTTCCAAAGAACTCCCGCGGC 330  
Db 946 S.S.N..B..C..WHY...G.N.DM.Y.H..MTB..T.BW.ABT.HMAHNB..THRGH.. 887  
Qy 331 GCTACTGCGGCGCAACATGGGCGATTTGCACTGTGTTGCTCGAGTGTCTTAATAGTTT 390  
Db 886 S..Y.RCM.MTMG..DR.RWBHM.R.DSTAKKBK...SBS.S.S...NA.H..GAR.H.S 827  
Qy 391 ATGCACACTCCCAATTAACACACAGGATGTTGGACAGACTTAAGGGCGCAAGT 450  
Db 826 .BS.CMDTC.CBGBBHT.R..AB.ADNBWB.BETHAG..YND..KM.NN.BN... 767  
Qy 451 GGCCTGAGCCGCCCACTTACCAACCAACGAGATGAGCCGCCCAATGTCACT 510  
Db 766 SAM..K..MNS.S.H.RY...TH...RMA.SY..NNM...NM--..SYGNS.Y..BAA..N 709  
Qy 511 TTGCGTATCGGCGCAACGCTGTTGACTGCTGTGCGACAGTCTCTTAATCAGGC 570  
Db 708 DS.H.A..NM.G..MMWS.SM.CNMT..S.NM.N..KCH.BCS...S.SNMS.H...M.. 649  
Qy 571 GCGGCGATACCCCTAAAGTTCCCGGATCCAGATACCGGTTCCAAACGACATCTC 630  
Db 648 NM...RYNHC.YW.BH.DNCT...AY.S.S.KAYSASS.HNHMK..M..BM.NMA.N. 589  
Qy 631 GAGAGTGAATCTCAATCACTGACGACCAAGCGGCGCAAGCTGTGATATGTT 690  
Db 588 KDA.C.S..RDSNS.H.B..MD...TASD..H.YNNS.H.RY..T....SCN.A.S. 529  
Qy 691 GCGCGTCCCGCGCGCGCGCGCTGCTTATGGCAGCGCTTTGACAGATCCGACTGA 750  
Db 528 .NCS.N.GM..S.D..HSSMB...SNMS.H.A.CMRGS.NBSK..ST.MNM.YGSYTRMH 469  
Qy 751 AGGATGCAATAGAGAGATCCCGCGGACAGTCAATAGACGCGGACCGAGTGTAT 810  
Db 468 RY.RD...B...RS.BSAYTNSG..CB.SSHCS.S.MMS.H.BM.YM..M.KS.NT.M. 409  
Qy 811 TCCTTGATATCATCTCAAGCACTTCTATGTTGATGTTGATGAGCTTGCGCTGAC 870  
Db 408 .NAC.H.B..N..ND..S.N.SM..M.BCY.Y..M.SM.SS.T.S.NKSTYRB...N. 349  
Qy 871 AATCAGAGCACTACCCAGAGTAGTTTGAAGCTCTCTTCAACAGCCAGAGAGCGG 930  
Db 348 S.H...S...D..MTHC..MT..N.S.H...NM.NM.WY..BD..SNSD.Y.BM..A. 289  
Qy 931 CCTGAGATGCGACGCTTTTGAAGCTGACACTTGATGAGTCTCTTCAAGAGAGA 990  
Db 288 TH.TNB...SNMG..TG...TNKGY.GNCS.H....N.NN.HH...MT.KNNR.R.. 229  
Qy 991 ACAACCAAGTCCACATAGATAC 1014  
Db 228 A.AS..BA.BTABNNHM.S.BM.M 205

RESULT 15  
 US-10-140-805-358/C  
 ; Sequence 358, Application US/10140805  
 ; Publication No. US20030207417A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Baker, Kevin P.  
 ; APPLICANT: Beresini, Maureen  
 ; APPLICANT: DeForge, Laura  
 ; APPLICANT: Desnoyers, Luc  
 ; APPLICANT: Filvaroff, Ellen  
 ; APPLICANT: Gao, Wei-Qiang  
 ; APPLICANT: Gerltsen, Mary E.  
 ; APPLICANT: Goddard, Audrey  
 ; APPLICANT: Godowski, Paul J.  
 ; APPLICANT: Gurney, Austin L.  
 ; APPLICANT: Sherwood, Steven  
 ; APPLICANT: Smith, Victoria  
 ; APPLICANT: Stewart, Timothy A.  
 ; APPLICANT: Thomas, Daniel  
 ; APPLICANT: Watanabe, Colin K  
 ; APPLICANT: Wood, William  
 ; APPLICANT: Zhang, Zemin  
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC  
 ; TITLE OF INVENTION: ACIDS ENCODING THE SAME  
 ; FILE REFERENCE: P33081C176  
 ; CURRENT APPLICATION NUMBER: US/10/140,805  
 ; CURRENT FILING DATE: 2002-05-07  
 ; Prior Application removed - See File Wrapper or Palm  
 ; NUMBER OF SEQ ID NOS: 550  
 ; SEQ ID NO 358  
 ; LENGTH: 1049  
 ; TYPE: PRT  
 ; ORGANISM: Homo Sapien  
 ; US-10-140-805-358

Query Match 1.1%, Score 44.8, DB 17, Length 1049;  
 Best Local Similarity 5.2%; Pred. No. 0.0049;  
 Matches 42; Conservative 229; Mismatches 531; Indels 2; Gaps 1;  
 QY 211 TCCTGTAACGAGGAGGAGCACTTCTTATCGAGCCAAAGCTGAGCTGCACTGG 270  
 Db 1006 YMY...MSM...M.....BHBH.HK.Y.DS.R.T.M.N.AMTRHAKB.BAMMS... 947  
 QY 271 GTGCCCAAGCGCCGACGCGCACTGACAGCGCTTCGTGTCAGGAACCTCCCGCGCC 330  
 Db 946 S.S.N..B..C..MHY..G.N.DM.Y.H.MTB..T.BW.ABT.HMAHRB..THRGH. 887  
 QY 331 GCTACTGCGCGCAACATGGGCACTGTCAGACTGTGTGCTCGAGTGTCTTAATAGTTT 390  
 Db 886 S..Y.RCM.MTMG.DR.RWBHM.R.DSTAKBK...SBS.S.S...NA.H.GAR.H.S 827  
 QY 391 ATGCCACTGCCATTAACACGCGAGTGCAGCGTTCGAGAACTCAAGCGGCCCAAGT 450  
 Db 826 .BS.CMDTTC.CBGBBHT.R..AB.ADNBWB.WBTHAG..YVDD...RM.NN.BN... 767  
 QY 451 GGCGTGAAGCGCCCAACCTTACCAACACCAACGCAAGATGAGCCCGCAATGTCAT 510  
 Db 766 SAM..K..MNS.H.RV...TH...RMA.SY..NNM..NM--.SYGNS.Y..BAA..N 709  
 QY 511 TTCGCTGATCGCCCAACGCTGCTGACTGCTCTGCGACAGTCTCTTATTACAGGC 570  
 Db 708 DS..H.A..NM.G..MMS.SM.CMNT.S.NM.N..KCH.BCS..S.SNMS.H...M.. 649  
 QY 571 GCGGCGATACCCCAAGTTCGCCGATCCAGTCAAGTACCGGTTCCAAACGACATCTC 630  
 Db 648 NM...RYNHC.YW.BH.DNCY...AY.S.S.KAYSASS.HNHNK.M..BM.NMA.N. 589  
 QY 631 GAGAGTGAATCTCAATCACTGACCAAGCGCGCCAAAGGCAAGCTGTCTGATAATGTT 690  
 Db 588 KDA.C.S..RDSNS.H.B..MD...TASD..H.YNNS.H.RY..T.....SCN.A.S. 529

QY 691 GCGCTGCGCGCGCGCGCGCGCGCTTCTTATTCGAGCGCTTTGACAAAGTACGACTCCA 750  
 Db 528 .NCS.N.GM..S.D..HSSMB...SNMS.H.A.CRMGS.NBSK..ST.MNM.YGSTRMH 469  
 QY 751 AGGCATGCCAATGAGAGATCCACGCGGACAGATAGACGAGCGGACGAGTGAT 810  
 Db 468 RY.RD...B..RS.BSATYNSG.CB.SSHCS.S.MBS.H.BM.YM..M.KS.NT.M. 409  
 QY 811 TCCTTGATACATCTCAAGCACTTCTTATGATGCTTTCAGCGTTCGCGCTCAC 870  
 Db 408 .NAC.H.B..N..ND..S.N.SM..M.BCY.Y..M.SM.SS.T.S.NKSTRB...N. 349  
 QY 871 AATCAGACACTACCAAGTACTTTGAGGCTCTCTTACACGCCAGAGAAAGCG 930  
 Db 348 S.H...S...D..MTHC..MT..N.S.H...MN.NM.WY..BD..SNSD.Y.BM..A. 289  
 QY 931 CCGTGATGCGCACGGCTTTGAGGCTGACACTGATGAGTCTCTGACAAAGAGA 990  
 Db 288 TH.TNB...SNMG.TG...TNRGY.GNCS.H.....N.NN.HH..MT.KNNR.R.. 229  
 QY 991 ACAACCAAGTCCACATAGATAAC 1014  
 Db 228 A.AS..BA.BTANNHM.S.BM.M 205

Search completed: August 11, 2005, 03:50:12  
 Job time : 10853 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM nucleic - nucleic search, using sw model

Run on: August 10, 2005, 17:59:01 ; Search time 12576 Seconds

(without alignments)  
12730.464 Million cell updates/sec

Title: US-09-913-878a-1\_COPY\_2447\_6652

Perfect score: 4206  
Sequence: 1 atgacaccattactactctag.....tcacagggaatgcgattat 4206Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

## Database :

EST:\*  
1: gb\_eest1:\*  
2: gb\_eest2:\*  
3: gb\_hnc:\*  
4: gb\_eest3:\*  
5: gb\_eest4:\*  
6: gb\_eest5:\*  
7: gb\_eest6:\*  
8: gb\_g881:\*  
9: gb\_g882:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	78	1.9	737	7	CF883391 trlC088xh
2	78	1.9	776	6	CB809520 trlC088xh
3	78	1.9	787	6	CB907399 trlC088xh
4	76.4	1.8	972	7	CF885460 trlC088xh
5	51.8	1.2	786	9	CG809201 FSAE20TF
6	49	1.2	570	2	BE433646 EST400175
7	49	1.2	684	7	CK851141 11217 Stc
8	49	1.2	737	7	CK854274 15379 Stc
9	47	1.1	603	7	CV190148 vC001-6ms
10	47	1.1	925	9	CNS0091P
11	45.6	1.1	597	2	AM649811 EST328265
12	45	1.1	367	2	BF112386 EST439976
13	45	1.1	421	1	AT896465 EST65896
14	45	1.1	525	2	AM933692 EST359535
15	44.8	1.1	742	7	CF820058 EST697440
16	44.6	1.1	681	8	BH045290 RPCI-24-2
17	44	1.0	310	1	AA501295 vH62d09.r
18	44	1.0	454	2	AA259012 z635901.s
19	44	1.0	529	1	AW971719 EST83808
20	44	1.0	584	5	BP380240 BP380240
21	44	1.0	592	2	AM872991 hq21a01.x
22	44	1.0	708	6	CA180649 SCACST315
23	44	1.0	720	9	AY409056 Homo sapi
24	44	1.0	720	9	AY409057 Pan trogl

25	44	1.0	958	1	AL559045
26	44	1.0	1844	3	CR609043 full1-Leng
27	44	1.0	3022	3	AL833977 Homo sapi
28	43.6	1.0	1011	1	AL559276
29	43.4	1.0	521	5	BQ111608 EST597184
30	43.4	1.0	610	2	BB920043 EST423812
31	42.6	1.0	787	9	AL098749 Drosoph11
32	42.6	1.0	1206	9	CL979698 OaIFCC033
33	42.4	1.0	582	4	BI246074 IPI 68 D1
34	42.4	1.0	584	4	BP379014 BP379014
35	42.4	1.0	622	5	BU965311 BAC08907.
36	42.4	1.0	641	6	CA257764 SCCCF110
37	42.4	1.0	681	1	A1125852 qe01f06.x
38	42.2	1.0	405	7	CN253957 BIP0019 B
39	41.8	1.0	966	9	CG006611 ZUAH588TV
40	41.2	1.0	555	6	CA866785 i472e02.x
41	41.2	1.0	563	5	BQ271751 i472e02.x
42	41.2	1.0	595	1	A1764981 wh57c04.x
43	41.2	1.0	651	5	BQ014248 UI-H-BD1-
44	41.2	1.0	658	8	B2336439 h234c12.b
45	41.2	1.0	715	7	CP865881 trlC003xm

## ALIGNMENTS

RESULT 1  
CF883391  
LOCUS CF883391 737 bp mRNA linear EST 31-OCT-2003  
DEFINITION trlC088xh09.b1 T.reesei mycelial culture, Version 6 October 2003  
ACCESSION HYPOCREA jecorina CDNA clone trlC088xh09, mRNA sequence.  
VERSION CF883391  
KEYWORDS CF883391.1 GI:38138073  
SOURCE EST.  
ORGANISM Hypocrea jecorina (anamorph: Trichoderma reesei)  
Hypocrea jecorina  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Hypocreomycetidae; Hypocerales; Hypocreaceae; Hypocrea.  
1 (bases 1 to 737)

REFERENCE  
AUTHORS Diener, S.B., Dunn-Coleman, N., Foreman, P., Houfek, T.D.,  
Teunissen, P.J.M., van Solingen, P., Dankmeyer, U., Mitchell, T.K.,  
Ward, M. and Dean, R.A.  
TITLE Characterization of the protein processing and secretion pathways  
in a comprehensive set of expressed sequence tags from Trichoderma  
reesei

JOURNAL  
COMMENT FEMS Microbiol. Lett. 230 (2), 275-282 (2004)  
Contact: Ralph A. Dean  
Fungal Genomics Laboratory  
North Carolina State University  
Campus Box 7251, Raleigh, NC 27695, USA  
Tel: 919-513-0020  
Fax: 919-513-0024  
Email: ralph.dean@ncsu.edu  
Seq primer: IT-F1 primer.  
Location/Qualifiers

FEATURES  
source  
1..737  
/organism="Hypocrea jecorina"  
/mol\_type="mRNA"  
/strain="QM6a"  
/db\_xref="taxon:51453"  
/clone="trlC088xh09"  
/dev\_stage="mycelia"  
/clone\_lib="T.reesei mycelial culture, Version 6 October 2003"  
/note="Vector: pREP3, Site\_1: Not I/Sal I; Mycelial culture grown from 24 hrs to 6 days with varying Carbon and Nitrogen sources and concentrations."

## ORIGIN

Query Match 1.9%; Score 78; DB 7; Length 737;  
Best Local Similarity 54.5%; Pred. No. 1.2e-10;  
Matches 156; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 3889 GAGATGATATCATGGCGATTGCTGAGGGCTAGACGGCTTTCAAGCTGTATCAACCAAG 3948  
 DB 221 GCGTTGAGTACTGTTTCTTTTCTGAAACATGCGACCACTTTTCAAGAGTATCAACCAAGC 280  
 QY 3949 AGCCCCAAGTTTGTGTGGCAGATGGCGGAGACAGCTCGGATACATTAAAGGCGCAGATG 4008  
 DB 281 GCGCCCAAGTTTGTGTGAAACATGCGGAAAGCAGCTTGTCTACATCAAAAGGCGATGCGC 340  
 QY 4009 ACAGACGAGCCCGGTGAAGCGCGCCCGCTTGAATGACCGCTTCATATGCGGCTTG 4068  
 DB 341 CGAAACGACGCTGGCGAGACGTCATAAGGTGCTGTGATCCCGAGATGTGGGGGCTCTG 400  
 QY 4069 ATGCGCGATTAAGAAGTTTACGAAGCATATGTGCGCCAGCTCGAGGCGCATGATCGGAG 4128  
 DB 401 CGGCGGACAAAGAGCTCATCACTCGCTGCGCATACAGAGGAGGCTGCGCGGATTCG 460  
 QY 4129 TACCCTGATCCGAGGCTATGAAGTGTGGGCGATGATTTTG 4174  
 DB 461 GAGAGTCACTGCTTGTGAGAGAGGTGTTGAGTTTGAACATTAATG 506

RESULT 2  
 CB909520 776 bp mRNA linear EST 02-JUL-2003  
 LOCUS tric088xh09 T. reesei mycelial culture, Version 3 april Hypocrea  
 DEFINITION jecorina cDNA clone tric088xh09, mRNA sequence.  
 ACCESSION CB909520  
 VERSION CB909520.1 GI:30124181  
 KEYWORDS EST.  
 SOURCE Hypocrea jecorina (anamorph: Trichoderma reesei)  
 ORGANISM Hypocrea jecorina

REFERENCE  
 AUTHORS Foreman, P.K., Brown, D.E., Dankmeyer, L., Dean, R., Diener, S.,  
 Dunn-Coleman, N.S., Goedegebuur, F., Houfek, T.D., England, G.J.,  
 Kelley, A.S., Meerman, H.J., Mitchell, T., Mitchinson, C.,  
 Olivares, H.A., Teunissen, P.J., Yao, J., and Ward, M.  
 Transcriptional regulation of biomass-degrading enzymes in the  
 filamentous fungus Trichoderma reesei  
 J. Biol. Chem. 278 (34), 31988-31997 (2003)

TITLE JOURNAL  
 MEDLINE 22803314  
 PUBMED 12788920  
 COMMENT Contact: Pamela K. Foreman  
 Genencor Intl.  
 925 Page Mill Road, Palo Alto, CA 94304, USA  
 Tel: (650) 846-7635  
 Fax: (650) 621-7817  
 Email: PForeman@genencor.com  
 Seq primer: LT-F1 primer.  
 Location/Qualifiers

FEATURES  
 source  
 1..776  
 /organism="Hypocrea jecorina"  
 /mol\_type="mRNA"  
 /strain="QM6a"  
 /db\_xref="taxon:51453"  
 /clone="tric088xh09"  
 /dev\_stage="mycelia"  
 /clone\_1lb="T. reesei mycelial culture, Version 3 april"  
 /note="Vector: pREP3y; Site 1: Not I/Sal I; Mycelial  
 culture grown from 24 hrs to 6 days with varying Carbon  
 and Nitrogen sources and concentrations."

ORIGIN  
 Query Match 1.9%; Score 78; DB 6; Length 776;  
 Best Local Similarity 54.5%; Pred. No. 1.2e-10;  
 Matches 156; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 3889 GAGATGATATCATGGCGATTGCTGAGGGCTAGACGGCTTTCAAGCTGTATCAACCAAG 3948  
 DB 221 GCGTTGAGTACTGTTTCTTTTCTGAAACATGCGACCACTTTTCAAGAGTATCAACCAAGC 280  
 QY 3949 AGCCCCAAGTTTGTGTGGCAGATGGCGGAGACAGCTCGGATACATTAAAGGCGCAGATG 4008  
 DB 281 GCGCCCAAGTTTGTGTGAAACATGCGGAAAGCAGCTTGTCTACATCAAAAGGCGATGCGC 340  
 QY 4009 ATGCGCGATTAAGAAGTTTACGAAGCATATGTGCGCCAGCTCGAGGCGCATGATCGGAG 4128  
 DB 401 CGGCGGACAAAGAGCTCATCACTCGCTGCGCATACAGAGGAGGCTGCGCGGATTCG 460

DB 281 GCGCCCAAGTTTGTGTGAAACATGCGCGAAAGCAGCTTGCCTTACATCAAGGAGATGCGC 340  
 QY 4009 ACAGACGAGACCCGGTGAAGCGCGCCCGGCTTATGACCGCTTCTATGTATGCGGCTTG 4068  
 DB 341 CGAAAGCAGCTGCGGAGACGTCATAAGGTGCTGTGATCCCGAGATGTGGGGGCTCTG 400  
 QY 4069 ATGCGCGATTAAGAAGTTTACGAAGCATATGTGCGCCAGCTCGAGGCGCATGATCGGAG 4128  
 DB 401 CGGCGGACAAAGAGCTCATCACTCGCTGCGCATACAGAGGAGGCTGCGCGGATTCG 460  
 QY 4129 TACCCTGATCCGAGGCTATGAAGTGTGGGCGATGATTTTG 4174  
 DB 461 GAGAGTCACTGCTTGTGAGAGAGGTGTTGAGTTTGAACATTAATG 506

RESULT 3  
 CB907399 787 bp mRNA linear EST 02-JUL-2003  
 LOCUS tric082xd19 T. reesei mycelial culture, Version 3 april Hypocrea  
 DEFINITION jecorina cDNA clone tric082xd19, mRNA sequence.  
 ACCESSION CB907399  
 VERSION CB907399.1 GI:30122057  
 KEYWORDS EST.  
 SOURCE Hypocrea jecorina (anamorph: Trichoderma reesei)  
 ORGANISM Hypocrea jecorina

REFERENCE  
 AUTHORS Foreman, P.K., Brown, D.E., Dankmeyer, L., Dean, R., Diener, S.,  
 Dunn-Coleman, N.S., Goedegebuur, F., Houfek, T.D., England, G.J.,  
 Kelley, A.S., Meerman, H.J., Mitchell, T., Mitchinson, C.,  
 Olivares, H.A., Teunissen, P.J., Yao, J., and Ward, M.  
 Transcriptional regulation of biomass-degrading enzymes in the  
 filamentous fungus Trichoderma reesei  
 J. Biol. Chem. 278 (34), 31988-31997 (2003)

TITLE JOURNAL  
 MEDLINE 22803314  
 PUBMED 12788920  
 COMMENT Contact: Pamela K. Foreman  
 Genencor Intl.  
 925 Page Mill Road, Palo Alto, CA 94304, USA  
 Tel: (650) 846-7635  
 Fax: (650) 621-7817  
 Email: PForeman@genencor.com  
 Seq primer: LT-F1 primer.  
 Location/Qualifiers

FEATURES  
 source  
 1..787  
 /organism="Hypocrea jecorina"  
 /mol\_type="mRNA"  
 /strain="QM6a"  
 /db\_xref="taxon:51453"  
 /clone="tric082xd19"  
 /dev\_stage="mycelia"  
 /clone\_1lb="T. reesei mycelial culture, Version 3 april"  
 /note="Vector: pREP3y; Site 1: Not I/Sal I; Mycelial  
 culture grown from 24 hrs to 6 days with varying Carbon  
 and Nitrogen sources and concentrations."

ORIGIN  
 Query Match 1.9%; Score 78; DB 6; Length 787;  
 Best Local Similarity 54.5%; Pred. No. 1.2e-10;  
 Matches 156; Conservative 0; Mismatches 130; Indels 0; Gaps 0;

QY 3889 GAGATGATATCATGGCGATTGCTGAGGGCTAGACGGCTTTCAAGCTGTATCAACCAAG 3948  
 DB 215 GCGTTGAGTACTGCGAGCTGCTGAAAGCTTTGACACATTTTCAAGAGTATCAACCAAGC 274  
 QY 3949 AGCCCCAAGTTTGTGTGGCAGATGGCGGAGACAGCTCGGATACATTAAAGGCGCAGATG 4008  
 DB 275 GCGCCCAAGTTTGTGTGAAACATGCGCGAAAGCAGCTTGTCTACATCAAGGCGATGCGC 334  
 QY 4009 ACAGACGAGACCCGGTGAAGCGCGCCCGGCTTATGACCGCTTCTATGTATGCGGCTTG 4068

Db 335 CGAAGCAGCGCTGGCGAGAGCTCAAGAGCTGTGATGCCGAGATGTGGGGCTCTG 394  
Qy 4069 ATGCCGATAGAGTTTACAGACAGATGTGGCCAGGCTGAGGGGAGATGATCGGAG 4128  
Db 395 CGGCGGAGCAAGAGCTATCACTCGCTGGCGATACAGAGGAGGCTGCCCGGAGATTGG 454  
Qy 4129 TACCTGATCCGAGGCTATGATGAGTGTGGCGATGATTTTG 4174  
Db 455 GAGAGTGCATGCTTGGAGAGGTGTGAGTTTACGATATG 500

RESULT 4  
CF885460 972 bp mRNA linear EST 31-OCT-2003  
LOCUS trico82xd19.b1.T.reesei mycelial culture, Version 6 October 2003  
DEFINITION Hypocrea jecorina cDNA clone trico82xd19, mRNA sequence.  
ACCESSION CF885460  
VERSION CF885460.1 GI:38140142  
KEYWORDS EST.  
SOURCE Hypocrea jecorina (anamorph: Trichoderma reesei)  
ORGANISM Hypocrea jecorina  
Bukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Hypocreomycetidae; Hypocreales; Hypocreaceae; Hypocrea.  
REFERENCE 1 (bases 1 to 972)  
Diener, S.E., Dunn-Coleman, N., Foreman, P., Houfek, T.D.,  
Teunissen, P.J.M., van Solingen, P., Dankmeyer, L., Mitchell, T.K.,  
Ward, M. and Dean, R.A.  
Characterization of the protein processing and secretion pathways  
in a comprehensive set of expressed sequence tags from Trichoderma  
reesei  
JOURNAL FEWS Microbiol. Lett. 230 (2), 275-282 (2004)  
COMMENT Contact: Ralph A. Dean  
Fungal Genomics Laboratory  
North Carolina State University  
Campus Box 7251, Raleigh, NC 27695, USA  
Tel: 919-513-0020  
Fax: 919-513-0024  
Email: ralph.dean@ncsu.edu  
Seq primer: LT-F1 primer  
Location/Qualifiers  
source 1..972  
/organism="Hypocrea jecorina"  
/mol\_type="mRNA"  
/strain="OM6a"  
/db\_xref="taxon:51453"  
/clone="trico82xd19"  
/dev\_strage="mycelia"  
/clone\_1lb="T. reesei mycelial culture, Version 6 October  
2003"  
/note="Vector: pREP3; Site 1: Not I/Sal I; Mycelial  
culture grown from 24 hrs to 6 days with varying carbon  
and Nitrogen sources and concentrations."

ORIGIN  
Query Match 1.8%; Score 76.4; DB 7; Length 972;  
Best Local Similarity 54.2%; Pred. No. 3.9e-10;  
Matches 155; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

Qy 4129 TACCTGATCCGAGGCTATGAGTGTGGCGATGATTTTG 4174  
Db 378 GAGAGTGCATGCTTGGAGAGGTGTGAGTTTACGATATG 423

RESULT 5  
CG809201/c 786 bp DNA linear GSS 13-NOV-2003  
LOCUS FSAAE20TR largeinsertGenomicLibrary Fusarium virguliforme genomic  
DEFINITION clone KMFV2C15, genomic survey sequence.  
ACCESSION CG809201  
VERSION CG809201.1 GI:38262675  
KEYWORDS GSS.  
SOURCE Fusarium virguliforme  
ORGANISM Fusarium virguliforme  
Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
Hypocreomycetidae; Hypocreales; Nectriaceae; Nectria.  
REFERENCE 1 (bases 1 to 786)  
Meksem, K., Ishihara, H., Koo, H., Shultz, J., Ali, S., Iqbal, J.,  
Lightfoot, D.A. and Town, C.D.  
End sequencing of BACs from a fingerprint physical map of the  
causative agent of soybean sudden death syndrome, Fusarium  
virguliforme  
JOURNAL Unpublished (2003)  
COMMENT Other GSSs: FSAAE20TR  
Contact: Chris Town and K. Meksem  
The Center of Excellence in Soybean Research, Teaching and  
Outreach, Southern Illinois University at Carbondale and Plant  
Genomics, The Institute for Genomic Research  
Room 176, Ag. Building, Mail Code 4415, Carbondale, IL 62901-4415,  
USA and 9712 Medical Center Drive, Rockville, MD 20850, USA  
Tel: 618 453 3103 and 301-838-3523  
Fax: 618 453-7457 and 301-838-0208  
Email: meksemk@uiu.edu; cdtown@tigr.org (URL:  
http://Fusariumvirguliforme.uiu.edu)  
Seq primer: CAGGAAACAGCTATGACC  
Class: BAC ends.  
Location/Qualifiers  
source 1..786  
/organism="Fusarium virguliforme"  
/mol\_type="genomic DNA"  
/cultivar="Monticello"  
/db\_xref="taxon:232082"  
/clone="KMFV2C15"  
/clone\_1lb="largeinsertGenomicLibrary"  
/note="Organ: Hyphae; Vector: pINDIOBAC5; A single spore  
derived culture was used. Hyphae were grown in an  
incubator for four days. Nuclei were isolated and embedded  
in agarose, restriction digested with Hind III. Large size  
DNA fragments were ligated in vector pINDIOBAC5 and  
electro-transformed into DH10B cells."

ORIGIN  
Query Match 1.2%; Score 51.8; DB 9; Length 786;  
Best Local Similarity 48.8%; Pred. No. 0.0058;  
Matches 140; Conservative 0; Mismatches 147; Indels 0; Gaps 0;

QY 4123 TCGAGTACCTGATCCGAGGTCTATGAAGTCTGGCCGATGTA 4169  
DB 352 GACATGAGAGTGTGACGCGCTCGAGAGGTGATGATGATGTA 306

RESULT 6  
BE433646  
LOCUS BE433646 570 bp mRNA linear EST 18-MAY-2001  
DEFINITION EST400175 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA  
ACCESSION BE433646  
VERSION BE433646  
KEYWORDS BE433646.1 GI:9431489  
SOURCE EST.  
ORGANISM Lycopersicon esculentum (tomato)  
Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 570)  
Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,  
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Rønning,C.M.,  
Nieman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and  
Tanksey,S.D.  
Generation of ESTs from tomato fruit tissue, breaker stage  
Unpublished (2000)  
CONTACT: CUGI  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.  
Location/Qualifiers  
1..570  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CLEG16J9"  
/tissue\_type="Pericarp"  
/dev\_stage="breaker"  
/lab\_host="SOLR"  
/clone\_lib="tomato breaker fruit, TIGR"  
/note="Vector: BluescriptSKmudapt; Site\_1: EcoRI;  
Site\_2: XhoI; Fruit were harvested at the breaker stage  
(first sign of lycopen accumulation on the blossom end  
of the fruit). Fruit were cut in half and the seeds and  
locules were discarded prior to freeing the pericarp."

ORIGIN  
Query Match 1.2%; Score 49; DB 2; Length 570;  
Best Local Similarity 51.6%; Pred. No. 0.034;  
Matches 112; Conservative 0; Mismatches 105; Indels 0; Gaps 0;

QY 2850 TACACTCTATCGACTGTATGTCCTCGTGCCTCCGACCCATTTCCCTAGTGA 2909  
DB 289 TAAATTATTTCTGAAGGAAATGTGTGTGCAAAAATCCATGCTTGATCCTGCTGA 348  
QY 2910 TATCCACGGGTTGACAGCTCTTCAAGCAGAGCTCCAGTCTCAAGATGTAATCAT 2969  
DB 349 TATTCGTTTAAAGCTGTAATATGTTGAGCGCTGCACACATGTAATGTGTGT 408  
QY 2970 CTCTCTACTAAAGAGATGATCGCTTGTCTAAGACCTATCTGTGTGAGACTACGACGG 3029  
DB 409 ATTCCCTCAGAAAGAAAGAACCTCATCCGAATGAATGTTCTGGAGTGAATTTGATGG 468  
QY 3030 CGAATAGCGCTGGTCTGTGTGGATCCGAGATGTC 3066  
DB 469 GGATATCTACTTGTGTTGTGTGGATCAAGACATGATC 505

RESULT 7  
CK851141

LOCUS CK851141 684 bp mRNA linear EST 08-MAR-2004  
DEFINITION 11217 Stolon Solanum tuberosum cDNA, mRNA sequence.  
ACCESSION CK851141  
VERSION CK851141.1 GI:45239751  
KEYWORDS EST.  
SOURCE Solanum tuberosum (potato)  
ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum.  
1 (bases 1 to 684)  
Flinn,B., Rothwell,C., Sardana,R., Griffiths,R., Lague,M., De  
koeijer,D., Audy,P., Goyer,C., Li,X.-Q., Wang-Pruski,G. and Regan,S.  
Generation of ESTs from stolon tissues of potato  
Unpublished (2004)  
CONTACT: Barry Flinn  
The Canadian Potato Genome Project - Biotechnology  
921 College Hill Rd, Fredericton, ON, E3B 6Z9, CANADA  
Email: [Bflinn@biotechnology.nb.ca](mailto:Bflinn@biotechnology.nb.ca)  
Seq primer: T3.  
Location/Qualifiers  
1..684  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Shepody"  
/db\_xref="taxon:4113"  
/tissue\_type="Stolon"  
/lab\_host="X110-Gold"  
/clone\_lib="Stolon"  
/note="Vector: Bluescript II SK(+) XR, Site\_1: EcoRI;  
Site\_2: XhoI; supplier: Developmental series. Plants from  
pathogen-free Solanum tuberosum var. Shepody, clone 1756,  
nuclear stock were grown in a greenhouse under natural  
conditions. RNA was isolated from stolon tissue."

ORIGIN  
Query Match 1.2%; Score 49; DB 7; Length 684;  
Best Local Similarity 53.4%; Pred. No. 0.036;  
Matches 103; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

QY 2870 ATGTCTCGGGGGGCGATCCCGACCCATTTCCCTAGTATATCCACGGGTTGCGACG 2929  
DB 171 ATGTGTTGTTGCAAAAATTCATGCTTCATCTGTAATTTCTGTTTAAAGGCTG 230  
QY 2930 TCTTCAAGCAGAGCTCCACAGTCTCAAGATGTAATCATCTTCTACTAAAGAGATG 2989  
DB 231 TAGATGTTCCAGGCTGCACACATGTAATGTTGTAATTTCCCGCAAGAAAGAAAA 290  
QY 2990 TACCGTTGCTAAGAGCTATCTGTGTGAGACTACGACCGCATATGCGCTGCTGCT 3049  
DB 291 GACCTCATCCGAATGAATGTTCTGGAGATGATGAGATATCTACTTTGTTGCT 350  
QY 3050 GGGATCCGAGAT 3062  
DB 351 GGGATCAAGACAT 363

RESULT 8  
CK854274  
LOCUS CK854274 737 bp mRNA linear EST 08-MAR-2004  
DEFINITION 15379 Stolon Solanum tuberosum cDNA, mRNA sequence.  
ACCESSION CK854274  
VERSION CK854274.1 GI:45242874  
KEYWORDS EST.  
SOURCE Solanum tuberosum (potato)  
ORGANISM Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum.  
1 (bases 1 to 737)  
Flinn,B., Rothwell,C., Sardana,R., Griffiths,R., Lague,M., De  
koeijer,D., Audy,P., Goyer,C., Li,X.-Q., Wang-Pruski,G. and Regan,S.  
Generation of ESTs from stolon tissues of potato

Full sequence and original trace file are available from the Plant Genome Network website (<http://pgn.cornell.edu>)

Full sequence and original trace file are available from the Plant  
Genome Network website (<http://pgn.cornell.edu>)  
Plate: vcc001-6ms4 row: C column: 07  
Seq primer: M13f.  
Location/Qualifiers  
1..603

```
/organism="Vaccinium corymbosum"  
/mol_type="mRNA"
```

```

/organism="Vaccinium corymbosum"
/mol_type="mRNA"
/db_xref="taxon:69266"
/clone="vc001-6m4-c07"
/tissue_type="flower buds with bud scales"
/lab_host="SOLR"
/clone_id="vc001"
/note="Vector: pBluescript SK (+/-); Site 1: EcoRI;
Site 2: XhoI; This is a directionally cloned,
non-normalized library. This library has been generated by
the Floral Genome Project (FGP). The Floral Genome Project
is funded by NSF's Plant Genome Research Program
(DBI-0115684). More information about the project can be
obtained at http://fgp.bio.psu.edu"

```

Query Match	1.1%;	Score 47;	DB 7;	Length 603;
Best Local Similarity	51.3%;	Pred. No. 0.13;		
Matches 137;	Conservative	0;	Mismatches 125;	Indels 5;
				Gaps 1;

2788 CTTGAGGAAATGAGGTTTCATGTCGGATTCTCTCAAGTTCAGAGACGAGGAGGAGTCT 2847

Db 138 CTCGTTGAAATTTGTTTGTGAATCATGGATCTAGGTTTTCAGAAACAAGAAGATCT 197

2848 TTACACTCCTATCGGACTGTGATGTCCTCGTGGCGGATCCCGCCATTCCCTAGT 2907

Db  
198 TCAAGTGATTAGGGGTTTGTGGT-----GTAGCGAAGATCCCTGTCTTCAACCAGGA 252

2908 GATATCCAACGGTTCCAGCAGTCTTCAAGCAGAGTCCACAGTCTCAAGGATGTAATC 2967

253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1047 1048 1049 1050 1051 1052 1053 1054 1055 1056 1057 1058 1059 1060 1061 1062 1063 1064 1065 1066 1067 1068 1069 1070 1071 1072 1073 1074 1075 1076 1077 1078 1079 1080 1081 1082 1083 1084 1085 1086 1087 1088 1089 1090 1091 1092 1093 1094 1095 1096 1097 1098 1099 1100 1101 1102 1103 1104 1105 1106 1107 1108 1109 1110 1111 1112 1113 1114 1115 1116 1117 1118 1119 1120 1121 1122 1123 1124 1125 1126 1127 1128 1129 1130 1131 1132 1133 1134 1135 1136 1137 1138 1139 1140 1141 1142 1143 1144 1145 1146 1147 1148 1149 1150 1151 1152 1153 1154 1155 1156 1157 1158 1159 1160 1161 1162 1163 1164 1165 1166 1167 1168 1169 1170 1171 1172 1173 1174 1175 1176 1177 1178 1179 1180 1181 1182 1183 1184 1185 1186 1187 1188 1189 1190 1191 1192 1193 1194 1195 1196 1197 1198 1199 1200 1201 1202 1203 1204 1205 1206 1207 1208 1209 1210 1211 1212 1213 1214 1215 1216 1217 1218 1219 1220 1

1  
 2  
 3  
 4  
 5  
 6  
 7  
 8  
 9  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525

\_\_\_\_\_

2000 2001 2002 2003 2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818

vy 3028 GCGAHMGGCIGGICIGCIGGAH 3034

Db 373 GGGGATTGTACTTTGTCACTTGGGAT 399

РЕШЕНИЕ 10

CNS0091P/c

DEFINITION *Drosophila melanogaster* genome survey sequence TET3 end of BAC #

fly), genomic survey sequence.

```

ACCESSION      AL053013
VERSION        AL053013.1  GI:4934461

```

KEYWORDS	GSS.
SOURCE	Protophila melanogaster (fruit fly)

ORGANISM Drosophila melanogaster  
Eukaryota: Metazoa: Arthropoda: Insecta: Diptera: Drosophila

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

REFERENCE 1 (bases 1 to 925)

*****	*****
TITLE	Direct Submission

JOURNAL  
submitted (02-JUN-1999) GENOSCOPE - Centre National de Séquençage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : [secref@genoscope.cns.fr](mailto:secref@genoscope.cns.fr))

COMMENT

- Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)

Determination of this RAC-end sequence was carried out as part of a

collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the *Drosophila*

melanogaster genome using these BACs. For further information

melanogaster BAC library was prepared by Kazutoyo Osogawa and



Aaron Mammoeser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's P1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at [http://bacpac.med.buffalo.edu/drosophila\\_bac.htm](http://bacpac.med.buffalo.edu/drosophila_bac.htm).

#### FEATURES

source  
Location/Qualifiers  
1..925  
/organism="Drosophila melanogaster"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:7227"  
/clone="BACRI9D16"  
/clone\_1lb="RPCI-98"  
/note="Tend : TET3"

#### ORIGIN

Query Match 1.1%; Score 47; DB 9; Length 925;  
Best Local Similarity 16.2%; Pred. No. 0.15;  
Matches 51; Conservative 142; Mismatches 119; Indels 3; Gaps 1;  
QY 32 GCCCGCTGAGGAATATAACCGGCTCAATAACGACTACCTGGCCTCCAGTGTG 91  
DB 918 SCSSBSBSSSSSSSTSSNSBSSBSSBSSSTSSBSSBSSSSSSGTTSSA 859  
QY 92 TCGGAGACACACTCTCACCCCGCCGGAAGAGAGTGGCCGAGTACGAGGATT 151  
DB 858 CVKGNASSCGCCCGGABCCMCSSSSCCGSABARVYRASGAGKRGSGSASASH 799  
QY 152 TCGGTCCCATGACAGATCTACAGAGCC--CTGAACTTCTCTACTGGCGGAAGATG 208  
DB 798 SSSBACBSSSSSCASCMASASSSSSRBSGCGAGSGASSRSSSSSASAGSV 739  
QY 209 ACTCCCTGACACGAGAGCCAACTTCTTCTGAGGCGCAAGCTCGGAGCTGCACT 268  
DB 738 SSASSSSSSSSSVSCSVASMSBSSBSASASSSSSSSCCCTCTSWSC 679  
QY 269 GGGTGCACCAAGCCGACGACCTGACGCTTCCTGGTCCAGAGAACTCCCGCG 328  
DB 678 STGASMEARSSSSSSSSSSSSMSASSASASSSSSSSGSAGBSMSSGGG 619  
QY 329 CCGCTACTGCCGCC 343  
DB 618 SGSVASSSGMSSSVS 604

RESULT 11  
AM649811 597 bp mRNA linear EST 18-MAY-2001  
LOCUS EST328265 tomato germinating seedlings, TAMU Lycopersicon  
DEFINITION esculentum cDNA clone cLEI9L3 5', mRNA sequence.  
ACCESSION AM649811  
VERSION AM649811.1 GI:7411049  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 597)  
REFERENCE Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,  
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Romning,C.M.,  
Niernan,W., Fraser,C.M., Giovannoni,J.J., Martin,G.B. and  
Tanksey,S.D.  
Generation of ESTs from germinating tomato seed  
TITLE Unpublished (2000)  
JOURNAL Contact: CUGI  
COMMENT Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>

5 prime sequence.  
Location/Qualifiers  
1..597  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cLEI9L3"  
/issue\_type="whole seedlings"  
/dev\_stage="7 days post imbibition"  
/clone\_1lb="tomato germinating seedlings, TAMU"  
/note="Vector: pBluescript SK(-); Site\_1: EcoRI; Site\_2:  
XhoI; 7 days post imbibition on water-agar. Mixed stage  
whole germinating seedlings from seed coat emergence up  
to two centimeters in seeds not showing obvious signs of  
germination were discarded."

#### ORIGIN

Query Match 1.1%; Score 45.6; DB 2; Length 597;  
Best Local Similarity 50.9%; Pred. No. 0.34;  
Matches 108; Conservative 0; Mismatches 104; Indels 0; Gaps 0;  
QY 2855 TCCTATCGACTGTATGTCCTCGTGGCGGAGATCCCAACCCATTCCTAGATATCC 2914  
DB 7 TCAATCTGAAGGAAATGGTGTGTGCAATATCAATGCTTCATCTCGATATGC 66  
QY 2915 AACGGTTCGACGAGCTTCAAGCCAGACCTCCACAGTCTCAAGATGTATCATCTCT 2974  
DB 67 GTGTTTAAAGCTGTAAATGTTTCGAGCGCTGCACACATGTAGATTGTGTATTC 126  
QY 2975 CTACTAAGAGATGTACCGCTTCTTAAGAGCTATCTGTGAGACTACGACGCGGATA 3034  
DB 127 CTCAGATGATGAAMAAACCTCTCCGAATGATTTCTGGAGTATTTGATGGAGATA 186  
QY 3035 TGGCTGGGTCTGTGGATTCGGAGATGCT 3066  
DB 187 TCTACTTTGTTTCTGGAGTCAAGACATGATC 218

RESULT 12  
BF112386 367 bp mRNA linear EST 18-MAY-2001  
LOCUS EST439976 tomato breaker fruit Lycopersicon esculentum cDNA clone  
DEFINITION CLEGA1011 5' sequence, mRNA sequence.  
ACCESSION BF112386  
VERSION BF112386.1 GI:10942076  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 367)  
REFERENCE Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,  
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Romning,C.M.,  
Niernan,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and  
Tanksey,S.D.  
Generation of ESTs from tomato fruit tissue, breaker stage  
TITLE Unpublished (2000)  
JOURNAL Contact: CUGI  
COMMENT Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>.

#### FEATURES

source  
Location/Qualifiers  
1..367  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="CLEGA1011"  
/issue\_type="Pericarp"  
/dev\_stage="Breaker"

/lab host="SOLR"  
/clone.lib="tomato breaker fruit"  
/note="Vector: pBluescriptSKmClnadpt; Site 1: EcorI;  
Site 2: XhoI; supplier: Boyce Thompson Institute;  
Sequencing: The Institute for Genomic Research. Fruit  
were harvested at the breaker stage (first sign of  
lycopene accumulation on the blossom end of fruit). Fruit  
were cut in half and the seeds and locules were discarded  
prior to freezing the pericarp."

## ORIGIN

Query Match 1.1%; Score 45; DB 2; Length 367;  
Best Local Similarity 54.5%; Pred. No. 0.44;  
Matches 90; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy 2902 CCTAGTATATCCAAAGCGGTTCCAGAGCTTTCAAGCCAGAGCTCCACAGTCTCAAGAT 2961  
Db 14 CCTGGTATATTCGTGTTTAAAGGCTGTAATGTTCCAGGCTGCACACATGATGAT 73  
Qy 2962 GTAATCATCTCTCTACTAAAGAGATGTACCGCTGCTAAGAGCTATCTGTGAGAC 3021  
Db 74 TGTGTATATTCCTCAGAAAGAAAAGACCTCATCCGAATGATGTTCTGGAGTGAT 133  
Qy 3022 TACGACGGCGATATGCGCTGGGCTGTCTGGGATCCGAGATCGTC 3066  
Db 134 TTGGATGGGATATCTACTTTGTTTCTGGGATCAAGCATGATC 178

RESULT 13  
LOCUS A1896465 421 bp mRNA linear EST 18-MAY-2001  
DEFINITION EST265896 tomato callus, TAMU Lycopersicon esculentum cDNA clone  
cLEJ515G23, mRNA sequence.  
ACCESSION A1896465  
VERSION A1896465.1 GI:5602355  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 421)  
Alcala,J., Vrebalov,J., White,R., Matera,A.L., Vision,T.,  
Holt,I.E., Liang,F., Upton,J., Craven,M.B., Bowman,C.L., Ann,S.,  
Ronnig,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and  
Giovannoni,J.  
Generation of ESTs from tomato callus tissue  
Unpublished (1999)  
Contact: CUGI  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.

REFERENCE  
AUTHORS

TITLE  
JOURNAL  
COMMENT

FEATURES  
SOURCE

1. 421  
Location/Qualifiers  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cLEJ515G23"  
/tissue\_type="callus"  
/dev\_stage="25-40 days old"  
/lab\_host="XLI-BIve MRP"  
/clone.lib="tomato callus, TAMU"  
/note="Vector: pBluescript SK(-); Site 1: EcorI; Site 2:  
XhoI; supplier: Giovannoni laboratory; cLEJ - Cotyledons  
of seedlings 7-10 days post-germination were excised, cut  
at both ends and placed on MS medium with no selection.  
Mixed callus was harvested at 25 and 40 days and included  
undifferentiated masses. Tomato Callus EST Library"

## ORIGIN

Query Match 1.1%; Score 45; DB 1; Length 421;  
Best Local Similarity 54.5%; Pred. No. 0.46;  
Matches 90; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy 2902 CCTAGTATATCCAAAGCGGTTCCAGAGCTTTCAAGCCAGAGCTCCACAGTCTCAAGAT 2961  
Db 14 CCTGGTATATTCGTGTTTAAAGGCTGTAATGTTCCAGGCTGCACACATGATGAT 73  
Qy 2962 GTAATCATCTCTCTACTAAAGAGATGTACCGCTGCTAAGAGCTATCTGTGAGAC 3021  
Db 74 TGTGTATATTCCTCAGAAAGAAAAGACCTCATCCGAATGATGTTCTGGAGTGAT 133  
Qy 3022 TACGACGGCGATATGCGCTGGGCTGTCTGGGATCCGAGATCGTC 3066  
Db 134 TTGGATGGGATATCTACTTTGTTTCTGGGATCAAGCATGATC 178

RESULT 14  
LOCUS AM933692 525 bp mRNA linear EST 18-MAY-2001  
DEFINITION EST359535 tomato fruit mature green, TAMU Lycopersicon esculentum  
cDNA clone cLEJ56C18 5', mRNA sequence.  
ACCESSION AM933692  
VERSION AM933692.1 GI:8109093  
KEYWORDS EST.  
SOURCE Lycopersicon esculentum (tomato)  
ORGANISM Lycopersicon esculentum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanales; Solanaceae; Solanum; Lycopersicon.  
1 (bases 1 to 525)  
Alcala,J., Vrebalov,J., White,R., Matera,A.L., Holt,I.E., Liang,F.,  
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ann,S.,  
Ronnig,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and  
Giovannoni,J.  
Generation of ESTs from tomato fruit tissue  
Unpublished (1999)  
Contact: CUGI  
Clemson University Genomics Institute  
100 Jordan Hall, Clemson, SC 29634, USA  
Email: <http://www.genome.clemson.edu/orders/index.html>  
5 prime sequence.

REFERENCE  
AUTHORS

TITLE  
JOURNAL  
COMMENT

FEATURES  
SOURCE

1. 525  
Location/Qualifiers  
/organism="Lycopersicon esculentum"  
/mol\_type="mRNA"  
/cultivar="TA496"  
/db\_xref="taxon:4081"  
/clone="cLEJ56C18"  
/tissue\_type="fruit pericarp"  
/dev\_stage="mature green (3-5 days pre-ripening)"  
/lab\_host="SOLR"  
/clone.lib="tomato fruit mature green, TAMU"  
/note="Vector: pBluescript SK(-); Site 1: EcorI; Site 2:  
XhoI; cLEJ - Fruit were tagged at the 1cm stage and  
harvested 3-5 days prior to ripening. Fruit were cut in  
half to verify the seeds were indeed 'immature' and the  
seeds and locules were discarded prior to freezing the  
pericarp"

## ORIGIN

Query Match 1.1%; Score 45; DB 2; Length 525;  
Best Local Similarity 54.5%; Pred. No. 0.49;  
Matches 90; Conservative 0; Mismatches 75; Indels 0; Gaps 0;

Qy 2902 CCTAGTATATCCAAAGCGGTTCCAGAGCTTTCAAGCCAGAGCTCCACAGTCTCAAGAT 2961  
Db 14 CCTGGTATATTCGTGTTTAAAGGCTGTAATGTTCCAGGCTGCACACATGATGAT 73  
Qy 2962 GTAATCATCTCTCTACTAAAGAGATGTACCGCTGCTAAGAGCTATCTGTGAGAC 3021  
Db 74 TGTGTATATTCCTCAGAAAGAAAAGACCTCATCCGAATGATGTTCTGGAGTGAT 133

QY 3022 TACGACGGGATATGCGCTGGCTCTGCGATCCGAGATCGTC 3066  
 Db 134 TTGGATGGGATATCTACTTTGTTTCTGGATCAAGATGATC 178

## RESULT 15

CF820058 742 bp mRNA linear EST 01-APR-2004  
 LOCUS EST697440 *Coccidioides posadasii* saprobic phase cDNA library, 2 to  
 DEFINITION 4 kb *Coccidioides posadasii* cDNA clone CIDA224 3' end, mRNA  
 sequence.

ACCESSION CF820058  
 VERSION CF820058.1 GI:45926115  
 KEYWORDS EST.  
 SOURCE *Coccidioides posadasii*  
 ORGANISM *Coccidioides posadasii*  
 Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
 Onygenales; mitosporic Onygenales; *Coccidioides*.

REFERENCE 1 (bases 1 to 742)  
 AUTHORS Gardner M.J. and Cole G.T.  
 TITLE Analysis of gene expression in *Coccidioides posadasii* mycelia and  
 spherules via expressed sequence tags

JOURNAL Unpublished (2003)  
 COMMENT Other ESTs: EST697441  
 Contact: Gardner MJ  
 The Institute for Genomic Research  
 9712 Medical Center Drive, Rockville, MD 20850, USA  
 Tel: 301 838 3519  
 Fax: 301 838 0208  
 Email: gardner@igf.org

FEATURES  
 source Location/Qualifiers  
 1..742  
 /organism="Coccidioides posadasii"  
 /mol\_type="mRNA"  
 /strain="C735"  
 /db\_xref="taxon:199306"  
 /clone="CIDA224"  
 /dev\_stage="saprobic phase (mycelia)"  
 /lab\_host="E. coli DH10B, T1 phage resistant"  
 /clone\_1lb="Coccidioides posadasii saprobic phase cDNA  
 library, 2 to 4 kb"  
 /note="Vector: pExpress 1; Site\_1: Not I; Site\_2: Eco RV;  
 Coccidioides posadasii saprobic phase cDNA library, size  
 fractionated cDNA 2 to 4 kb"

## ORIGIN

Query Match 1.1%; Score 44.8; DB 7; Length 742;  
 Best Local Similarity 52.1%; Pred. No. 0.63;  
 Matches 100; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 2874 CCTGTGCGCGATGCCAGCCATTCCTAGTATATCAACGGGTTGCAGCAGTCTT 2933  
 Db 46 CATATTTGCCGAATCGTCTTGATCCGGCGACATCCGGTCTTAAGTGTCAA 105  
 QY 2934 CAAGCCAGAGCTCCAGTCTCAAGATGTAATCATCTTCTACTAAGAGATGACC 2993  
 Db 106 TGCACCGCGCTTACCATCTTAGAGATGTATGTATCCACAGACAGAGATAGGA 165  
 QY 2994 GCTTGTAAAGACTATCTGTGAGACTACAGCGGATATAGCCTGGTCTGTGGGA 3053  
 Db 166 TATCCCAAGCATGTTGGGTGGTGTGACTTGATGTGATGATTAACCTGTCAATTGGGA 225  
 QY 3054 TCCGAGATCGT 3065  
 Db 226 CCAAGATCTCCT 237

Search completed: August 11, 2005, 04:59:29  
 Job time : 12585 secs